INDIAN AGRICULTURAL
RESEARCH INSTITUTE, NEW DELHI.

I. A. R. I. 6.
MGIPC—S4—10 AR—21-6-49—1,000.

PROCEEDINGS

OF THE

ROYAL SOCIETY OF LONDON

SERIES A. MATHEMATICAL AND PHYSICAL SCIENCES

VOL 180

LONDON
Printed and published for the Royal Society
By the Cambridge University Press
Bentley House, N.W.1

3 July 1942

PRINTED IN GREAT BRITAIN BY
WALTER LEWIS, M.A.
AT THE CAMBRIDGE UNIVERSITY PRESS

CONTENTS

SERIES A VOL 180

No. A 980. 18 March 1942

PAGE

Bakerian Lecture. The physical interpretation of quantum mechanics. By P. A. M. Dirac, F.R.S. 1

Molecular structure and rubber-like elasticity. I. The crystal structures of β gutta-percha, rubber and polychloroprene. By C. W. Bunn . 40

Molecular structure and rubber-like elasticity. II. The stereochemistry of chain polymers. By C. W. Bunn 67

Molecular structure and rubber-like elasticity. III. Molecular movements in rubber-like polymers. By C. W. Bunn 82

The structure of polychloroprene. By C. J. B. Clews. (Plate 1) . . 100

The lattice spacings of the primary solid solutions in magnesium of the metals of group III B and of tin and lead. By G. V. Raynor . . 107

No. A 981. 5 June 1942

Further studies on active nitrogen. III. Experiments to show that traces of oxygen or other impurity affect primarily the walls of the vessel, and not the phenomena in the gas space. By Lord Rayleigh, F.R.S. 123

Further studies on active nitrogen. IV. The ionisation associated with active nitrogen. By Lord Rayleigh, F.R.S. 140

Surface energy relations in liquid/solid systems. I. The adhesion of liquids to solids and a new method of determining the surface tension of liquids. By C. Ockrent and G. Macdougall 151

Stress systems in aeolotropic plates. IV. By A. E. Green 173

The effect of a free surface on compressional (shock) waves. By W. G. Bickley 209

The 'second maximum' of the shower transition curve of cosmic radiation. By E. P. George, L. Jánossy and M. McCaig 219

The distribution of autelectronic emission from single crystal metal points. II. The adsorption, migration and evaporation of thorium, barium and sodium on tungsten and molybdenum. By M. Benjamin and R. O. Jenkins. (Plates 2–11) 225

The photosensitivity of visual purple solutions and the scotopic sensitivity of the eye in the ultra-violet. By C. F. Goodeve, F.R.S., R. J. Lythgoe and E. E. Schneider. (Abstract) 236

No. A 982. 5 June 1942

PAGE

The detection and inhibition of free radical chain reactions. By R. E. Smith and C. N. Hinshelwood, F.R.S. 237

On the significance tests for the introduction of new functions to represent measures. By H. Jeffreys, F.R.S. 256

Imperfections in the structure of cobalt. I. Experimental work and proposed structure. By O. S. Edwards and H. Lipson. (Plate 12) . 268

Imperfections in the structure of cobalt. II. Mathematical treatment of proposed structure. By A. J. C. Wilson 277

On the theory of finite deformations of elastic crystals. By R. Fürth . 285

The effect of thermal vibrations on the scattering of X-rays. II. By K. Sarginson 305

Cross-section measurements for disintegrations produced by deuterons in the heavy elements. By R. S. Krishnan and E. A. Nahum . . 321

Excitation function measurements for disintegrations produced by deuterons in the heavy elements. By R. S. Krishnan and E. A. Nahum 333

Arrangements with given number of neighbours. By T. S. Chang and C. C. Ho 345

No. A 983. 3 July 1942

Some practically important stress-systems in solids of revolution. By R. V. Southwell, F.R.S. 367

Effect of thermal vibrations on the scattering of X-rays. III. By M. Born, F.R.S. 397

Thermodynamics of a relativistic Fermi-Dirac gas. By D. S. Kothari and B. N. Singh 414

The photocombination of hydrogen and chlorine in oxygen-free systems. By M. Ritchie and D. Taylor 423

Crystal theory of metals: calculation of the elastic constants. By K. Fuchs and H. W. Peng 451

Diffuse scattering of X-rays by crystals. II. Detailed calculation of the surfaces of isodiffusion for the (002), (112), (222) and (110) reflexions of sodium single crystals. By H. A. Jahn 476

The free streaming of gases in sloping galleries. By E. H. M. Georgeson 484

Index 495

BAKERIAN LECTURE

The physical interpretation of quantum mechanics

By P. A. M. Dirac, F.R.S.

(*Delivered* 19 *June* 1941.—*Received* 23 *September* 1941)

Modern developments of atomic theory have required alterations in some of the most fundamental physical ideas. This has resulted in its being usually easier to discover the equations that describe some particular phenomenon than just how the equations are to be interpreted. The quantum mechanics of Heisenberg and Schrödinger was first worked out for a number of simple examples, from which a general mathematical scheme was constructed, and afterwards people were led to the general physical principles governing the interpretation, such as the superposition of states and the indeterminacy principle. In this way a satisfactory non-relativistic quantum mechanics was established.

In extending the theory to make it relativistic, the developments needed in the mathematical scheme are easily worked out, but difficulties arise in the interpretation. If one keeps to the same basis of interpretation as in the non-relativistic theory, one finds that particles have states of negative kinetic energy as well as their usual states of positive energy, and, further, for particles whose spin is an integral number of quanta, there is the added difficulty that states of negative energy occur with a negative probability.

With electrons the negative-probability difficulty does not arise, and one can get a sensible interpretation of the negative-energy states by assuming them to be nearly all occupied and an unoccupied one to be a positron. This model, however, is excessively complicated to work with and one cannot get any results from it without making very crude approximations. The simple accurate calculations that one *can* make would apply to a world which is almost saturated with positrons, and it appears to be a better method of interpretation to make the general assumption that transition probabilities obtained from these calculations for this hypothetical world are the same as those for the actual world.

With photons one can get over the negative-energy difficulty by considering the states of positive and negative energy to be associated with the emission and absorption of a photon respectively, instead of, as previously, with the existence of a photon. The simplest way of developing the theory would make it apply to a hypothetical world in which the initial probability of certain states is negative, but transition probabilities calculated for this hypothetical world are found to be always positive, and it is again reasonable to assume that these transition probabilities are the same as those for the actual world.

NON-RELATIVISTIC THEORY

Heisenberg, with his discovery of quantum mechanics, introduced a new outlook on the nature of physical theory. Previously, it was always considered essential that there should be a detailed description of what is taking place in natural phenomena, and one used this description to calculate results comparable with experiment. Heisenberg put forward the view that it is sufficient to have a mathematical scheme from which one can calculate in a consistent manner the results of all experiments—a detailed description in the traditional sense is unnecessary and may very well be impossible to establish.

Heisenberg's method focuses attention on to the quantities which enter into experimental results. It was first applied to the theory of spectra, for which these quantities are the energy levels of the atomic system and certain probability coefficients which determine the probability of a radiative transition taking place from one level to another. The method sets up equations connecting these quantities and allows one to calculate them, but does not go beyond this. It does not provide any description of radiative transition processes. It does not even allow one to deduce how the results of a calculation are to be used, but requires one to assume Einstein's laws of radiation (the laws which tell how the probability of a radiative transition process depends on the intensity of the incident radiation), and to assume that certain quantities determined by the calculation are the coefficients appearing in the laws.

Shortly after Heisenberg's discovery, Schrödinger set up independently another form of quantum mechanics, which also enables one to calculate energy levels and probability coefficients and gives results agreeing with those of Heisenberg, but which introduces an important new feature. *It connects together, in one calculation, a set of probability coefficients which act together under certain conditions in Nature*; for example, the set of probability coefficients referring to transitions from one particular initial state to any final state. In this respect it is to be contrasted with Heisenberg's method, which connects together in one calculation all the probability coefficients for a dynamical system, i.e. the probability coefficients from all initial states to all final states.

This feature of Schrödinger's method gives it two important advantages. First, as a consequence of its enabling one to obtain fewer results at a time, it makes the computation much simpler. Secondly, it supplies, in a certain sense, a *description* of what is taking place in Nature, since a calculation leading to results which come into play together under certain conditions in Nature will be in close correspondence with the physical process that is taking place under those conditions, various points in the calculation having their counterparts in the physical process. A description in this limited sense seems to be the most that is possible for atomic processes. It implies a much less complete connexion between the mathematics and the physics than one has in classical mechanics, and one might be disinclined to call it a description at all, but one may at least consider it as an appropriate generalization of what one usually means by a description. On account of Schrödinger's method allowing a description in this new sense while Heisenberg's allows none, Schrödinger's method introduces an outlook on the nature of physical theory intermediate between Heisenberg's and the old classical one.

When Heisenberg's & Schrödinger's theories were developed it was soon found that they both rested on the same mathematical formalism and differed only with regard to the method of physical interpretation. The formalism is a generalization of the Hamiltonian form of classical dynamics, involving linear operators instead of ordinary algebraic variables, and is so natural and beautiful as to make one feel sure of its correctness as the foundation of the theory. The question of its interpretation, however, which involved unifying Heisenberg's & Schrödinger's ideas into a satisfactory comprehensive scheme, was not so easily settled, and is probably still not finally settled.

The situation of a formalism becoming established before one is clear about its interpretation should not be considered as surprising, but as a natural consequence of the drastic alterations which the development of physics has required in some of the basic physical concepts. This makes it an easier matter to discover the mathematical formalism needed for a fundamental physical theory than its interpretation, since the number of things one has to choose between in discovering the formalism is very limited, the number of fundamental ideas in pure mathematics being not very great, while with the interpretation most unexpected things may turn up.

The best way of seeking the interpretation in such cases is probably from a discussion of simple examples. This way was used for the theory of quantum mechanics and led eventually to a satisfactory interpretation applicable to all phenomena for which relativistic effects are negligible. This interpretation is more closely connected with Schrödinger's method than Heisenberg's, as one would expect on account of the former affording in some sense a description of Nature, and is centred round a *Schrödinger's wave function*, which is one of the things that can be operated on by the linear operators which the dynamical variables have become. The correspondence which the existence of a description implies between the mathematics and the physics makes a wave function correspond to a state of motion of the atomic system, in such a way that, for example, a calculation *which gives the transition probabilities from a particular initial state to* any final state would be based on that wave function which represents the motion ensuing from this initial state. A wave function is a complex function $\psi(q_1 q_2 \dots q_n t)$ of all the co-ordinates $q_1, q_2, \dots, q_n$ of the system and of the time t, and it receives the interpretation that the square of its modulus, $|\psi(q_1 \dots q_n t)|^2$, is the probability, for the state of motion it corresponds to, of the co-ordinates having values in the neighbourhood of $q_1, q_2, \dots, q_n$, per unit volume of co-ordinate space, at the time t.

A wave function can be transformed so as to refer to other dynamical variables, for example, the momenta $p_1, p_2, \dots, p_n$, when it is said to be in *another representation*. The square of its modulus $|\psi(p_1 \dots p_n t)|^2$ is then the probability, per unit volume of momentum space, of the momenta having values in the neighbourhood of $p_1, p_2, \dots, p_n$ at the time t. A wave function itself never has an interpretation, but only the square of its modulus, and the need for distinguishing between two wave functions having the same squares of their moduli arises only because, if they are transformed to a different representation, the squares of their moduli will in general become different. This brings out the incompleteness of the kind of description which is possible with quantum mechanics.

One may make a slight modification in the wave functions in any representation by introducing a weight factor λ and arranging for the probability to be $\lambda|\psi|^2$ instead of $|\psi|^2$. The weight factor may be any positive function of the variables occurring in the wave function.

Wave functions have to satisfy a certain *wave equation*, namely, the equation

$$i\hbar\, d\psi/dt = H\psi, \tag{1}$$

where H is a Hermitian (or self-adjoint) linear operator, and is the Hamiltonian of the system expressed in the representation concerned. The wave equation (1) is a generalization of the Hamilton-Jacobi equation of classical mechanics. If S is a solution of the latter equation, then

$$\psi = e^{iS/\hbar} \tag{2}$$

will give a first approximation to a solution of the former.

An important property of the wave equation (1) is that it yields the conservation law—the total probability of the variables occurring in the wave function having any value is constant. The wave function should be normalized so as to make this probability initially unity and then it always remains unity. This conservation law is a mathematical consequence of the wave equation being linear in the operator d/dt and of H being a self-adjoint operator.

The wave equation is linear and homogeneous in the wave function ψ, and so are the transformation equations. In consequence, one can add together two ψ's and get a third. The correspondence between ψ's and states of motion now allows one to infer that there is a relationship between the states of motion, such that one can add or superpose two states to get a third. This relationship constitutes the Principle of Superposition of States, one of the general principles governing the interpretation of quantum mechanics that we are here considering.

Another of these principles is Heisenberg's Principle of Indeterminacy. This is a consequence of the transformation laws connecting $\psi(q)$ and $\psi(p)$, which show that each of these functions is the Fourier transform of the other, apart from numerical coefficients, so that one meets the same limitations in giving values to a q and p as in giving values to the position and frequency of a train of waves.

These general principles serve to bring out the departures needed from ordinary classical ideas. They are of so drastic and unexpected a nature that it is not to be wondered at that they were discovered only indirectly, as consequences of a previously established mathematical scheme, instead of being built up directly from experimental facts.

DIFFICULTIES OF THE RELATIVISTIC THEORY

The theory outlined above is not in agreement with the restricted principle of relativity, as is at once evident from the special role played by the time t. Thus, while it works very well in the non-relativistic region of low velocities, where it appears to be in complete agreement with experiment, it can be considered only as an approximation, and one must face the task of extending it to make it conform to restricted relativity. (General relativity need not be considered, since gravitational effects are negligible in atomic theory.) One should be prepared for possible further alterations being needed in basic physical concepts, and hence one should follow the route of first setting up the mathematical formalism and then seeking its physical interpretation.

Setting up the mathematical formalism is a fairly straightforward matter. One must first put classical mechanics into relativistic Hamiltonian form. One must take into account that the various particles comprising the dynamical system interact through the medium of the electromagnetic field, and one must use Lorentz's equations of motion for them, including the damping terms which express the reaction of radiation. This is done in Appendix I, where, with the help of an action principle, the equations of motion are obtained in the Hamiltonian form (1·40) with the Hamiltonians F_i, one for each particle, given by (1·39). This Hamiltonian formulation may now be made into a quantum theory by following rules which have become standardized from the non-relativistic quantum mechanics. The resulting formalism appears to be quite satisfactory mathematically, but when one proceeds to consider its physical interpretation one meets with serious difficulties.

Take an elementary example, that of a free particle without spin, moving in the absence of any field. The classical Hamiltonian for this system is the left-hand side of the equation

$$p_0^2 - p_1^2 - p_2^2 - p_3^2 - m^2 = 0, \qquad (3)$$

where p_0 is the energy and p_1, p_2, p_3 the momentum of the particle, the velocity of light being taken as unity. Passing over to quantum theory by the standard rules, one gets from this Hamiltonian the wave equation

$$(\hbar^2 \Box + m^2)\psi = 0, \qquad (4)$$

where
$$\Box \equiv \frac{\partial^2}{\partial x_0^2} - \frac{\partial^2}{\partial x_1^2} - \frac{\partial^2}{\partial x_2^2} - \frac{\partial^2}{\partial x_3^2}.$$

The wave function ψ here is a scalar, involving the co-ordinates x_1, x_2, x_3 and the time x_0 on the same footing, and so it is suitable for a relativistic theory.

If one now tries to use the old interpretation that $|\psi|^2$ is the probability per unit volume of the particle being in the neighbourhood of the point x_1, x_2, x_3 at the time x_0, one immediately gets into conflict with relativity, since this probability ought to transform under Lorentz transformations like the time component of a 4-vector, while $|\psi|^2$ is a scalar. Also the conservation law for total probability would no longer hold, the usual proof of it failing on account of the wave equation (4) not being linear in $\partial/\partial x_0$.

An important step forward was taken by Gordon (1926) and Klein (1927), who proposed that instead of $|\psi|^2$ one should use the expression

$$\frac{1}{4\pi i}\left\{\frac{\partial\bar{\psi}}{\partial x_0}\psi - \bar{\psi}\frac{\partial\psi}{\partial x_0}\right\}. \qquad (5)$$

This expression is the time component of a 4-vector. Further, it is easily verified that the divergence of this 4-vector vanishes, which gives the conservation law in relativistic form. Thus (5) is evidently the correct mathematical form to use.

This form leads to trouble on the physical side, however, since, although it is real, it is not positive definite like $|\psi|^2$. Its employment would result in one having at times a negative probability for the particle being in a certain place.

This is not the only difficulty. Let us consider the energy and momentum of the particle, and take for simplicity a state for which these variables have definite values. The corresponding wave function will be of the form of plane waves,

$$\psi = e^{-i(p_0x_0-p_1x_1-p_2x_2-p_3x_3)/h}.$$

In order that the wave equation (4) may be satisfied, the energy and momentum values p_0, p_1, p_2, p_3 here must satisfy the classical equation (3). This equation allows of negative values for the energy p_0 as well as positive ones and is, in fact, symmetrical between positive and negative energies. The negative energies occur also in the classical theory, but do not then cause trouble, since a particle started off in a positive-energy state can never make a transition to a negative-energy one. In the quantum theory, however, such transitions are possible and do in general take place under the action of perturbing forces.

The wave function may be transformed to the momentum and energy variables. The Gordon-Klein expression (5) then goes over into

$$|\psi(p_0p_1p_2p_3)|^2p_0^{-1}dp_1dp_2dp_3, \tag{6}$$

as the probability of the momentum having a value within the small domain $dp_1dp_2dp_3$ about the value p_1, p_2, p_3, with the energy having the value p_0, which must be connected with p_1, p_2, p_3 by (3). The weight factor p_0^{-1} appears in (6) and makes it Lorentz invariant, since $\psi(p)$ is a scalar—it is defined in terms of $\psi(x)$ to make it so—and the differential element $p_0^{-1}dp_1dp_2dp_3$ is Lorentz invariant. This weight factor may be positive or negative, and makes the probability positive or negative accordingly. Thus the two undesirable things, negative energy and negative probability, always occur together.

Let us pass on to another simple example, that of a free particle with spin half a quantum. The wave equation is of the same form (4) as before, but the wave function ψ is no longer a scalar. It must have two components, or four if there is a field present, and the way they transform under Lorentz transformations is given by the general connexion between the theory of angular momentum in quantum mechanics and group theory. The expression $\Sigma\,|\psi(x)|^2$, summed for the components of ψ, turns out to be the time component of a 4-vector, and further the divergence of this 4-vector vanishes. Thus it is satisfactory to use this expression as the probability per unit volume of the particle being at any place at any time. One does not now have any negative probabilities in the theory. However, the negative energies remain, as in the case of no spin.

We may go on and consider particles of higher spin. The general result is that there are always states of negative energy as well as those of positive energy. For particles whose spin is an integral number of quanta, the negative-energy states occur with a negative probability and the positive-energy ones with a positive probability, while for particles whose spin is a half-odd integral number of quanta, all states occur with a positive probability.

Negative energies and probabilities should not be considered as nonsense. They are well-defined concepts mathematically, like a negative sum of money, since the equations which express the important properties of energies and probabilities can still be used when they are negative. Thus negative energies and probabilities should be considered simply as things which do not appear in experimental results. The physical interpretation of relativistic quantum mechanics that one gets by a natural development of the non-relativistic theory involves these things and is thus in contradiction with experiment. We therefore have to consider ways of modifying or supplementing this interpretation.

PARTICLES OF HALF-ODD INTEGRAL SPIN

Let us first consider particles with a half-odd integral spin, for which there is only the negative-energy difficulty to be removed. The chief particle of this kind for which a relativistic theory is needed is the electron, with spin half a quantum. Now electrons, and also, it is believed, all particles with a half-odd integral spin, satisfy the exclusion principle of Pauli, according to which not more than one of them can be in any quantum state. (This principle is obtained in quantum mechanics from the requirement that wave functions shall be antisymmetric in all the particles.) With this principle there are only two alternatives for a state, either it is unoccupied or it is occupied by one particle, and a symmetry appears with respect to these two alternatives.

Some time ago I proposed a way of dealing with the negative-energy difficulty for electrons, based on a theory in which nearly all their negative-energy states are occupied. An unoccupied negative-energy state now appears as a 'hole' in the distribution of occupied negative-energy states and thus has a deficiency of negative energy, i.e. a positive energy. From the wave equation one finds that a hole moves in the way one would expect a positively charged electron to move. It becomes reasonable to identify the holes with the recently discovered positrons, and thus to get an interpretation of the theory involving positrons together with electrons. An electron jumping from a positive- to a negative-energy state in the theory is now interpreted as an annihilation of an electron and a positron, and one jumping from a negative- to a positive-energy state as a creation of an electron and a positron.

The theory involves an infinite density of electrons everywhere. It becomes necessary to assume that the distribution of electrons for which all positive-energy states are unoccupied and all negative-energy states occupied—what one may call the vacuum distribution, as it corresponds to the absence of all electrons and positrons in the interpretation—is completely unobservable. Only departures from this distribution are observable and contribute to the electric density and current which give rise to electromagnetic field in accordance with Maxwell's equations.

The above theory does provide a way out from the negative-energy difficulty, but it is not altogether satisfactory. The infinite number of electrons that it involves requires one to deal with wave functions of very great complexity and leads to such complicated mathematics that one cannot solve even the simplest problems accurately, but must resort to crude and unreliable approximations. Such a theory is a most inconvenient one to have to work with, and on general philosophical grounds one feels that it must be wrong.

Let us see whether one can modify the theory so as to make it possible to work out simple examples accurately, while retaining the basic idea of identifying unoccupied negative-energy states with positrons. The simple calculations that one can make involve simple wave functions, referring to only one or two electrons, and thus referring to nearly all the negative-energy states being unoccupied. The calculations therefore apply to a world almost saturated with positrons, i.e. having nearly every quantum state for a positron occupied. Such a world, of course, differs very much from the actual world. One can now calculate the probability of any kind of collision process occurring in this hypothetical world (in so far as electrons and positrons are concerned). One can deduce the probability coefficient for the process, i.e. the probability per unit number of incident particles or per unit intensity of the beam of incident particles, for each of the various kinds of incident particle taking part in the process. For this purpose one must use the laws of statistical mechanics, which tell how the probability of a collision process depends on the number of incident particles, paying due attention to the modified form of these laws arising from the exclusion principle of Pauli.

Let us now assume that *probability coefficients so calculated for the hypothetical world are the same as those of the actual world.* This single assumption provides a general physical interpretation for the formalism, enabling one to calculate collision probabilities in the actual world. It does not provide a complete physical theory, since it enables one to calculate only those experimental results that are reducible to collision probabilities, and some branches of physics, e.g. the structure of solids, do not seem to be so reducible. However, collision probabilities are the things for which a relativistic theory is at present most needed, and one may hope in the future to find ways of extending the scope of the theory to make it include the whole of physics.

Comparing the new theory with the old, one may say that the new assumption, identifying collision probability coefficients in the actual world with those in a certain hypothetical world, replaces the old assumption about the non-observability of the vacuum distribution of negative-energy electrons. The approximations needed for working out simple examples in the old theory are equivalent in their mathematical effect to making the new assumption; e.g. these approximations include the neglect of the Coulomb interaction between electron and positron in the calculation of the probability of pair creation and annihilation, and this interaction cannot appear in the new theory, since the calculation there is concerned with a one-electron system. Thus the new theory may be considered as a precise formulation of the old theory together with some general approximations needed for applying it.

The new theory for dealing with the negative-energy states of the electron may be applied to any kind of elementary particle with spin half a quantum, and probably also to particles with other half-odd integral spin values, provided, of course, they satisfy Pauli's exclusion principle. It may thus be applied to protons and neutrons. It requires for each particle the possibility of existence of an antiparticle of the opposite charge, if the original particle is charged. If the original particle is uncharged, one can arrange for the antiparticle to be identical with the original.

PARTICLES OF INTEGRAL SPIN

Most of the elementary particles of physics have half-odd integral spin, but there is the important exception of the light-quantum or photon, with spin one quantum, and there is the new cosmic-ray particle, the meson, also probably with spin one quantum, and other such particles may be discovered in the future. All these kinds of particle, it is believed, satisfy the Bose statistics, a statistics which allows any number of particles to be in the same quantum state with the same a priori probability. (This statistics is obtained in quantum mechanics from the requirement that wave functions shall be symmetric in all the particles.) For these kinds of particles the previous method of dealing with the negative-energy states is therefore no longer applicable, and there is the further difficulty of the negative probabilities.

When dealing with particles satisfying the Bose statistics, it is useful to consider the operators corresponding to the absorption of a particle from a given state or the emission into a given state. These operators can be treated as dynamical variables, although they do not have any analogues in classical mechanics. If one works out their equations of motion and transformation equations, one finds a remarkable correspondence. The absorption operators from a set of independent states have the same equations of motion and transformation equations as the wave function ψ representing a single particle, and similarly for the emission operators and the conjugate complex wave function $\bar{\psi}$. Thus one can pass from a one-particle theory to a many-particle theory by making the ψ and $\bar{\psi}$ describing the one particle into absorption and emission operators, which must satisfy the appropriate commutation relations. Such a passage is called second quantization.

One can get over the difficulties of negative energies and negative probabilities for Bose particles by abandoning the attempt to get a satisfactory theory of a single particle and passing on to consider the problem of many particles, using a method given by Pauli & Weisskopf (1934) for electrons having no spin and satisfying the Bose statistics. (Such electrons are not known experimentally, but there is no known theoretical reason why they should not exist.) The method of Pauli & Wiesskopf is to work entirely with positive-energy states. The operators of absorption from and emission into negative-energy states, arising in the application of second quantization to the one-electron theory, are replaced by the operators of emission into and absorption from positive-energy states of electrons with the opposite charge, respectively. This replacement does not disturb the laws of conservation of charge, energy and momentum. The resulting theory involves spinless electrons of both kinds of charge together, and leads to pair creation and annihilation, as with ordinary electrons and positrons.

The method of Pauli & Wiesskopf may be applied in a degenerate form to photons and leads to the quantum electrodynamics of Heisenberg & Pauli (1929*a* and *b*). To take into account that photons have no charge, one must start with a one-particle theory in which the wave functions are real, so that $\bar{\psi} = \psi$. The part of ψ referring to positive-energy states is then made into the absorption operators from positive-energy states, and the part referring to negative-energy states into the emission operators into positive-energy states. The resulting scheme of operators, involving only positive-energy photon states, may then be put into correspondence with classical electrodynamics, according to the usual laws governing the correspondence between quantum and classical theory.

It would seem that in this way the difficulties of negative energies and probabilities for Bose particles can be overcome, but a new difficulty appears. When one tries to solve the wave equation (or the wave equations if there are several particles with their respective Hamiltonians) one gets divergent integrals in the solution, of the form, in the case of photons,

$$\int_0^\infty f(\nu)\,d\nu, \quad f(\nu) \sim \nu^n \text{ for large } \nu, \tag{7}$$

ν being the frequency of a photon. The values 1, 0 and -1 for n are the chief ones occurring in simple examples. Thus the wave equation really has no solutions and the method fails.

I have made a detailed study of the divergent integrals occurring in quantum electrodynamics and have shown (Dirac 1939*b*) that all those with even values of n can be eliminated by introducing into the equations a certain limiting process, which one can justify by showing that a corresponding limiting process is needed in classical electrodynamics to get the equations of motion into Hamiltonian form. (It appears accordingly in the action principle of Appendix I.) The divergent integrals with odd values of n remain, however, and indicate something more fundamentally wrong with the theory.

Divergent integrals are a general feature of quantum field theories, and it has usually been supposed that they should be avoided by altering the forces or the laws of interaction between the elementary particles at small distances, so as to get the integrals cut off for some high value of ν. However, one can easily see that this is wrong, in the case of electrodynamics at any rate, by referring to the corresponding classical theory. The wave function should have its analogue in the solution of the Hamilton-Jacobi equation, in accordance with equation (2), but already when one tries to solve the Hamilton-Jacobi equation of classical electrodynamics corresponding to the wave equation of Heisenberg & Pauli's quantum electrodynamics, one meets with divergent integrals. Now the classical equations of motion concerned, namely, Lorentz's equations including radiation damping, have definite solutions when treated by straightforward methods and if, on trying to get these solutions by a Hamilton-Jacobi method, one meets with divergent integrals, it means simply that the Hamilton-Jacobi method is an unsuitable one, and not that one should try to alter the physical laws of interaction to get the integrals to converge. The correspondence between the quantum and classical theories is so close that one can infer that the corresponding divergent integrals in the quantum theory must also be due to an unsuitable mathematical method.

The appearance of divergent integrals with odd n-values in Heisenberg & Pauli's form of quantum electrodynamics may be ascribed to the unsymmetrical treatment of positive- and negative-energy photon states. If instead of using Pauli & Weisskopf's method one keeps to plain second quantization, one can build up a form of quantum electrodynamics symmetrical between positive- and negative-energy photon states, as is done in Appendix II. The new theory leads to similar equations as the old one, but with integrals of the type

$$\int_{-\infty}^{\infty} f(\nu)\, d\nu \tag{8}$$

instead of (7), and since $f(\nu)$ is always a rational algebraic function, and it is reasonable on physical grounds to approach the upper and lower limits of integration in (8) at the same rate, the divergencies with odd n-values all cancel out.

The work of Appendices I and II shows that the new form of quantum electrodynamics also corresponds to classical electrodynamics in accordance with the usual laws, with the exception that operators corresponding to real dynamical variables in the classical theory are no longer always self-adjoint. This exception is not important, as it rather stands apart from the general mathematical connexion between quantum and classical theory. The Hamilton-Jacobi equation corresponding to the wave equation of the new quantum electrodynamics differs from that of the old one only through being expressed in terms of a different set of co-ordinates, but the new Hamilton-Jacobi equation can be solved without divergent integrals and is connected with a satisfactory action principle, namely, that of Appendix I. Thus the correspondence with classical theory of the new form of quantum electrodynamics is more far-reaching than that of the old form, which provides a strong reason for preferring the new form. It now becomes necessary to find some new physical interpretation to avoid the difficulties of negative energies and probabilities.

Let us consider in more detail the relation between the two forms of quantum electrodynamics. In either form the electromagnetic potentials **A** at two points $\mathbf{x}'$ and $\mathbf{x}''$ must satisfy the commutation relations

$$[A_\mu(\mathbf{x}'), A_\nu(\mathbf{x}'')] = g_{\mu\nu}\Delta(\mathbf{x}'-\mathbf{x}''), \tag{9}$$

obtained from analogy with the classical theory, Δ being the four-dimensional Lorentz-invariant function introduced by Jordan & Pauli (1928), which has a singularity on the light-cone and vanishes everywhere else. In the quantum electrodynamics of Heisenberg and Pauli the **A**'s are operators referring to the absorption and emission of photons into positive-energy states. Let us call such operators $\mathbf{A}^1$. One could introduce a similar set of operators referring to the absorption and emission of photons into negative-energy states. Let us call these operators $\mathbf{A}^2$. They satisfy the same commutation relations (9) and commute with the $\mathbf{A}^1$'s. One can now introduce a third set of operators

$$\mathbf{A}^3 = 2^{-\frac{1}{2}}(\mathbf{A}^1+\mathbf{A}^2),$$

which operate on wave functions referring to photons in both positive- and negative-energy states, and which satisfy the same commutation relations (9). The use of this third set gives the new form of quantum electrodynamics arising from plain second quantization.

The three sets of **A**'s may be expressed in terms of their Fourier components thus:

$$\left.\begin{aligned} \mathbf{A}^1(\mathbf{x}) &= \iiint \{\mathbf{R}_\mathbf{k} e^{i(\mathbf{k},\mathbf{x})} + \overline{\mathbf{R}}_\mathbf{k} e^{-i(\mathbf{k},\mathbf{x})}\} k_0^{-1} dk_1 dk_2 dk_3, \\ \text{where}\qquad k_0 &= \sqrt{(k_1^2+k_2^2+k_3^2)}, \end{aligned}\right\} \tag{10}$$

$\mathbf{R}_\mathbf{k}$ being the emission operator and $\overline{\mathbf{R}}_\mathbf{k}$ the absorption operator,

$$\left.\begin{aligned} \mathbf{A}^2(\mathbf{x}) &= \iiint \{\mathbf{R}_\mathbf{k} e^{i(\mathbf{k},\mathbf{x})} + \overline{\mathbf{R}}_\mathbf{k} e^{-i(\mathbf{k},\mathbf{x})}\} k_0^{-1} dk_1 dk_2 dk_3, \\ \text{where}\qquad k_0 &= -\sqrt{(k_1^2+k_2^2+k_3^2)}, \end{aligned}\right\} \tag{11}$$

$$\mathbf{A}^3(\mathbf{x}) = 2^{-\frac{1}{2}}\Sigma_{k_0=\pm\sqrt{(k_1^2+k_2^2+k_3^2)}} \iiint \{\mathbf{R}_\mathbf{k} e^{i(\mathbf{k},\mathbf{x})} + \overline{\mathbf{R}}_\mathbf{k} e^{-i(\mathbf{k},\mathbf{x})}\} k_0^{-1} dk_1 dk_2 dk_3. \tag{12}$$

Since the three sets of **A**'s all satisfy the same commutation relations, they must correspond merely to three different representations of the same dynamical variables, and the passage from one to another must be a transformation of the linear type usual in quantum mechanics. Thus, after obtaining the divergency-free solution of the wave equation in the representation corresponding to $\mathbf{A}^3$, one could apply a transformation to get the solution in the $\mathbf{A}^1$ representation. However, the transformation would then introduce the same divergent integrals as appear with the direct solution of the wave equation in the $\mathbf{A}^1$ representation, so one would not get any further this way.

In working with the $\mathbf{A}^3$ representation one has redundant dynamical variables. It is as though, in dealing with a system of one degree of freedom with the variables q, p, one decided to treat it as a system of two degrees of freedom by putting

$$q = 2^{-\frac{1}{2}}(q_1+q_2), \quad p = 2^{-\frac{1}{2}}(p_1+p_2).$$

This would be quite a correct procedure, but would introduce an unnecessary complication. In the case of quantum electrodynamics, the complication is a necessary one, to avoid the divergent integrals. Let us put

$$\mathscr{B}(\mathbf{x}) = 2^{-\frac{1}{2}}\{\mathbf{A}^1(\mathbf{x})-\mathbf{A}^2(\mathbf{x})\}. \tag{13}$$

Then the $\mathscr{B}$'s commute with the $\mathbf{A}^3$'s, and thus with all the dynamical variables appearing in the Hamiltonian, so they are the redundant variables.

To determine the significance of redundant variables in quantum mechanics one may consider a general case and work in a representation which separates the redundant variables from the non-redundant ones. One then sees immediately that a solution of the wave equation corresponds in general, not to a single state, but to a set of states with a certain probability for each—what in the classical theory is called a Gibbs ensemble. The probabilities of the various states depend on the weights attached to the various eigenvalues of the redundant variables in the wave function, these weights being arbitrary, depending on the weight factor in the representation used. If one works in a representation which does not separate the redundant and non-redundant variables, as is the case in quantum electrodynamics with the representation corresponding to the use of $\mathbf{A}^3$, the general result that wave functions represent Gibbs ensembles and not single states must still be valid. Thus one can conclude that there are no solutions of the wave equation of quantum electrodynamics representing single states, but only solutions representing Gibbs ensembles. The problem remains of interpreting the negative energies and probabilities occurring with these Gibbs ensembles.

For any $\mathbf{x}$, $\mathscr{B}(\mathbf{x})$ commutes with the Hamiltonian and is a constant of the motion. We may give it any value we like, subject to not contradicting the commutation relations. Instead of $\mathscr{B}(\mathbf{x})$ it is more convenient to work with the potential field, $\mathbf{B}(\mathbf{x})$ say, obtained from $\mathscr{B}(\mathbf{x})$ by changing the sign of all the Fourier components containing $e^{ik_0x_0}$ with negative values of k_0. From (13), (10) and (11),

$$\mathbf{B}(\mathbf{x}) = 2^{-\frac{1}{2}}\Sigma\iiint\{\mathbf{R}_{\mathbf{k}}e^{i(\mathbf{k},\mathbf{x})}-\overline{\mathbf{R}}_{\mathbf{k}}e^{-i(\mathbf{k},\mathbf{x})}\}\,k_0^{-1}dk_1dk_2dk_3, \tag{14}$$

the Σ meaning a summation over both values of k_0, as in equation (12).

Let us now take $\mathbf{B}$ equal to the initial value of $\mathbf{A}^3$, a proceeding which does not contradict the commutation relations since its consequences are self-consistent. Then for the initial wave function ψ,

$$\{\mathbf{B}(\mathbf{x})-\mathbf{A}^3(\mathbf{x})\}\,\psi = 0.$$

or, from (12) and (14),

$$\bar{\mathbf{R}}_{\mathbf{k}}\,\psi = 0, \tag{15}$$

with k_0 either positive or negative. Thus any absorption operator applied to the initial wave function gives the result zero, which means that the corresponding state is one with no photons present.

The following natural interpretation for the wave function at some later time now appears. That part of it corresponding to no photons present may be supposed to give (through the square of its modulus) the probability of *no* change having taken place in the field of photons; that part corresponding to one positive-energy photon present may be supposed to give the probability of a photon having been emitted; that corresponding to one negative-energy photon present may be supposed to give the probability of a photon having been absorbed; and so on for the parts corresponding to two or more photons present. *The various parts of the wave function which referred to the existence of positive- and negative-energy photons in the old interpretation now refer to the emissions and absorptions of photons.* This disposes of the negative-energy difficulty in a satisfactory way, conforming to the laws of conservation of energy and momentum. It is possible only because of the redundant variables enabling one to arrange that the initial wave function shall correspond in its entirety to no emissions or absorptions having taken place.

The interpretation is not yet complete, because the theory at present would give a negative probability for a process involving the absorption of a photon, or the absorption of any odd number of photons. To find the origin of these negative probabilities, one must study the probability distribution of the photons initially present in the Gibbs ensemble, which one can do by transforming to the representation corresponding to the A^1 potentials. It is true that one cannot apply this transformation to a solution of the wave equation without getting divergent integrals, as has already been mentioned, but one can apply it to the initial wave function, which is of a specially simple form in the photon variables. This is done in Appendix III, and it is found that the probability of there being n photons initially in any photon state is $P_n = \pm 2$, according to whether n is even or odd. Strictly, to make $\Sigma_{n=0}^{\infty} P_n$ converge to the limit unity, one must consider P_n as a limit,

$$P_n = 2(\epsilon-1)^n, \tag{16}$$

with ϵ a small positive quantity tending to zero.

Probabilities 2 and -2 are, of course, not physically understandable, but one can use them mathematically in accordance with the rules for working with a Gibbs ensemble. One can suppose a hypothetical mathematical world with the initial probability distribution (16) for the photons, and one can work out the probabilities of radiative transition processes occurring in this world. One can deduce the corresponding probability coefficients, i.e. the probabilities per unit intensity of each beam of incident radiation concerned, by using Einstein's laws of radiation. For example, for a process involving the absorption of a photon, if the probability coefficient is B, the probability of the process is

$$\Sigma_{n=0}^{\infty} nP_n B = -\tfrac{1}{2}B, \tag{17}$$

and for a process involving the emission of a photon, if the probability coefficient is A, the probability of the process is

$$\Sigma_{n=0}^{\infty}(n+1)P_n A = \tfrac{1}{2}A. \tag{18}$$

Now the probability of an absorption process, as calculated from the theory, is negative, and that for an emission process is positive, so that, equating these calculated probabilities to (17) and (18) respectively, one obtains positive values for both B and A. Generally, it is easily verified that any radiative transition probability coefficient obtained by this method is positive.

It now becomes reasonable to assume that *these probability coefficients obtained for a hypothetical world are the same as those of the actual world.* One gets in this way a general physical interpretation for the quantum theory of photons. When applied to elementary examples, it gives the same results as Heisenberg & Pauli's quantum electrodynamics with neglect of the divergent integrals, since the extra factor $2^{-\frac{1}{2}}$ occurring in the matrix elements of the present theory owing to the $2^{-\frac{1}{2}}$ in the right-hand side of (12) compensates the factor $\frac{1}{2}$ in the right-hand side of (17) or (18). The present general method of physical interpretation is probably applicable to any kind of particle with an integral spin.

CONCLUSION

It appears that, whether one is dealing with particles of integral spin or of half-odd integral spin, one is led to a similar conclusion, namely, that the mathematical methods at present in use in quantum mechanics are capable of direct interpretation only in terms of a hypothetical world differing very markedly from the actual one. These mathematical methods can be made into a physical theory by the assumption that results about collision processes are the same for the hypothetical world as the actual one. One thus gets back to Heisenberg's view about physical theory—that all it does is to provide a consistent means of calculating experimental results. The limited kind of description of Nature which Schrödinger's method provides in the non-relativistic case is possible relativistically only for the hypothetical world, and even then is rather more indefinite (e.g. the principle of superposition of states no longer applies), because of the need to use a Gibbs ensemble for describing the photon distribution.

To have a description of Nature is philosophically satisfying, though not logically necessary, and it is somewhat strange that the attempt to get such a description should meet with a partial success, namely, in the non-relativistic domain, but yet should fail completely in the later development. It seems to suggest that the present mathematical methods are not final. Any improvement in them would have to be of a very drastic character, because the source of all the trouble, the symmetry between positive and negative energies arising from the association of energies with the Fourier components of functions of the time, is a fundamental feature of them.

APPENDIX 1. THE ACTION PRINCIPLE OF CLASSICAL ELECTRODYNAMICS

There are various forms which the action principle of classical electrodynamics may take, but most of them involve awkward conditions concerning the singularities of the field where the charged particles are situated and are not suitable for a subsequent passage to quantum mechanics. Fokker (1929) set up a form of action principle which does not refer to the singularities of the field and which appears to be the best starting point for getting a quantum theory. Fokker's action integral may conveniently be written with the help of the δ function and is then

$$S = S_1 + S_2,$$

where
$$S_1 = \Sigma_i m_i \int ds_i \tag{1·1}$$

and
$$S_2 = \Sigma_i \Sigma_{j \neq i} e_i e_j \iint \delta(\mathbf{z}_i - \mathbf{z}_j)^2 (\mathbf{v}_i, \mathbf{v}_j)\, ds_i\, ds_j. \tag{1·2}$$

The scalar product notation

$$(\mathbf{a}, \mathbf{b}) = a^\mu b_\mu = a_0 b_0 - a_1 b_1 - a_2 b_2 - a_3 b_3$$

is here used, and m_i and e_i are the mass and charge of the ith particle, the 4-vector $\mathbf{z}_i$ gives the four co-ordinates of the point on the world-line of the

ith particle whose proper-time is s_i, and $\mathbf{v}_i$ is the velocity 4-vector of the ith particle

$$\mathbf{v}_i = d\mathbf{z}_i/ds_i \tag{1·3}$$

satisfying

$$\mathbf{v}_i^2 = 1. \tag{1·4}$$

The integrals are taken along the world-lines of the particles, and the occurrence of the δ function $\delta(\mathbf{z}_i - \mathbf{z}_j)^2$ in S_2 ensures that the only values for $\mathbf{z}_i$ and $\mathbf{z}_j$ contributing to the double integral are those for which $(\mathbf{z}_i - \mathbf{z}_j)^2 = 0$, which means that each of the points $\mathbf{z}_i$, $\mathbf{z}_j$ is on the past or future light-cone from the other.

The action integral as it stands is not a general one covering all possible states of motion. To make it general one must, as has been pointed out by the author (1938), add to it a term of the form

$$S_3 = \Sigma_i e_i \int M_\mu(\mathbf{z}_i)\, v_i^\mu ds_i. \tag{1·5}$$

The 4-vector potential $M_\mu(\mathbf{x})$ may be left for the present an arbitrary function of the field point $\mathbf{x}$.

For the purpose of deducing the equations of motion one may take the limits of integration in the various integrals to be $-\infty$ and ∞, as was done by Fokker, but in order to introduce momenta and get the equations into Hamiltonian form one must take finite limits. Let us therefore suppose each s_i goes from s_i^0 to s_i', and let the corresponding $\mathbf{z}_i$ and $\mathbf{v}_i$ be $\mathbf{z}_i^0$, $\mathbf{z}_i'$ and $\mathbf{v}_i^0$, $\mathbf{v}_i'$. It is desirable to restrict the initial values s_i^0 so that the points $\mathbf{z}_i^0$ all lie outside each other's light-cones, and similarly with the final values s_i'. Thus

$$(\mathbf{z}_i^0 - \mathbf{z}_j^0)^2 < 0, \quad (\mathbf{z}_i' - \mathbf{z}_j')^2 < 0 \quad (i \neq j). \tag{1·6}$$

Before making variations in S, one should replace S_1 by

$$S_1' = \Sigma_i m_i \int \sqrt{\mathbf{v}_i^2} ds_i, \tag{1·7}$$

so as to make S homogeneous of degree zero in the differential elements ds_i, $\mathbf{v}_i$ counting as being of degree -1. The expression for S is then valid with s_i any parameter on the world-line of the ith particle, so that $\mathbf{v}_i$ defined by (1·3) does not necessarily satisfy (1·4).

Let us make variations $\partial\mathbf{z}_i(s_i)$ in the world-lines of the particles, $\partial\mathbf{M}(\mathbf{x})$ in the field function $\mathbf{M}(\mathbf{x})$, and Ds_i' in the final values of the s_i, so that the end-points of the world-lines are changed by

$$D\mathbf{z}_i' = \partial\mathbf{z}_i' + \mathbf{v}_i' Ds_i', \tag{1·8}$$

$\partial\mathbf{z}_i'$ being written for $\partial\mathbf{z}_i(s_i')$. The initial values of the s_i and the initial points of the world-lines we shall suppose for simplicity to be fixed, since variations in them would give rise to terms of the same form as those arising from variations in the final values and would not lead to anything new.

Varying S_1' given by (1·7) and using (1·4) after the variation process, one gets

$$\begin{aligned}\partial S_1' &= \Sigma_i m_i \left\{ \int_{s_i^0}^{s_i'} (\mathbf{v}_i, \partial \mathbf{v}_i)\, ds_i + Ds_i' \right\} \\ &= \Sigma_i m_i \left\{ -\int_{s_i^0}^{s_i'} (d\mathbf{v}_i/ds_i, \partial \mathbf{z}_i)\, ds_i + (\mathbf{v}_i', \partial \mathbf{z}_i') + Ds_i' \right\} \\ &= \Sigma_i m_i \left\{ -\int_{s_i^0}^{s_i'} (d\mathbf{v}_i/ds_i, \partial \mathbf{z}_i)\, ds_i + (\mathbf{v}_i', D\mathbf{z}_i') \right\} \qquad (1\cdot 9)\end{aligned}$$

with the help of (1·8). To get the variation in S_2 given by (1·2) one may, owing to the symmetry between i and j in the double sum, vary only quantities involving i and multiply by 2. The result is

$$\begin{aligned}\partial S_2 = \Sigma_i \Sigma_{j \neq i} e_i e_j &\left\{ \int_{s_i^0}^{s_i'} \int_{s_j^0}^{s_j'} \left[\frac{\partial \delta(\mathbf{z}_i - \mathbf{z}_j)^2}{\partial z_{\mu i}} \partial z_{\mu i} (\mathbf{v}_i, \mathbf{v}_j) + \delta(\mathbf{z}_i - \mathbf{z}_j)^2 (\mathbf{v}_j, \partial \mathbf{v}_i) \right] ds_i\, ds_j \right. \\ &\qquad \left. + \int_{s_j^0}^{s_j'} \delta(\mathbf{z}_i' - \mathbf{z}_j)^2 (\mathbf{v}_i', \mathbf{v}_j)\, Ds_i'\, ds_j \right\} \\ = \Sigma_i \Sigma_{j \neq i} e_i e_j &\left\{ \int_{s_i^0}^{s_i'} \int_{s_j^0}^{s_j'} \left[\frac{\partial \delta(\mathbf{z}_i - \mathbf{z}_j)^2}{\partial z_{\mu i}} (\mathbf{v}_i, \mathbf{v}_j) - \frac{d}{ds_i} \{\delta(\mathbf{z}_i - \mathbf{z}_j)^2 v_j^\mu\} \right] \partial z_{\mu i}\, ds_i\, ds_j \right. \\ &\qquad \left. + \int_{s_j^0}^{s_j'} \delta(\mathbf{z}_i' - \mathbf{z}_j)^2 (\mathbf{v}_j, D\mathbf{z}_i')\, ds_j \right\}. \qquad (1\cdot 10)\end{aligned}$$

In varying S_3 given by (1·5), one has to take into account that the total variation in $\mathbf{M}$ at a point $\mathbf{z}_i(s_i)$ on the ith world-line, let us call it $D\mathbf{M}(\mathbf{z}_i)$, consists of two parts, a part $\partial \mathbf{M}(\mathbf{z}_i)$ arising from the variation in the function $\mathbf{M}(\mathbf{x})$ and equal to the value of $\partial \mathbf{M}(\mathbf{x})$ at the point $\mathbf{x} = \mathbf{z}_i$, and a part arising from the variation in $\mathbf{z}_i$, equal to $\partial \mathbf{M}/\partial x_\nu$ at the point $\mathbf{x} = \mathbf{z}_i$ multiplied into $\partial z_{\nu i}$; thus

$$D\mathbf{M}(\mathbf{z}_i) = \partial \mathbf{M}(\mathbf{z}_i) + (\partial \mathbf{M}/\partial x_\nu)_{\mathbf{z}_i} \partial z_{\nu i}. \qquad (1\cdot 11)$$

The variation in S_3 is now

$$\begin{aligned}\partial S_3 &= \Sigma_i e_i \left\{ \int_{s_i^0}^{s'} [DM^\mu(\mathbf{z}_i)\, v_{\mu i} + M^\mu(\mathbf{z}_i)\, \partial v_{\mu i}]\, ds_i + M^\mu(\mathbf{z}_i')\, v_{\mu i}'\, Ds_i' \right\} \\ &= \Sigma_i e_i \left\{ \int_{s_i^0}^{s_i'} \left[\partial M^\mu(\mathbf{z}_i)\, v_{\mu i} + \left(\frac{\partial M^\mu}{\partial x_\nu} \right)_{\mathbf{z}_i} v_{\mu i}\, \partial z_{\nu i} - \frac{dM^\mu(\mathbf{z}_i)}{ds_i} \partial z_{\mu i} \right] ds_i + M^\mu(\mathbf{z}_i')\, Dz_{\mu i}' \right\}.\end{aligned}$$

(1·12)

The total variation in S is given by the sum of the three expressions (1·9), (1·10) and (1·12).

By equating to zero the total coefficient of $\partial z_{\mu i}$, one gets the equation of motion of the ith particle. It is

$$-m_i\frac{dv_i^\mu}{ds_i}+e_i\Sigma_{j\neq i}e_j\int_{s_j^0}^{s_j'}\left[\frac{\partial\delta(\mathbf{z}_i-\mathbf{z}_j)^2}{\partial z_{\mu i}}(\mathbf{v}_i,\mathbf{v}_j)-\frac{d}{ds_i}\{\delta(\mathbf{z}_i-\mathbf{z}_j)^2v_j^\mu\}\right]ds_j$$
$$+e_i\left[\left(\frac{\partial M^\nu}{\partial x_\mu}\right)_{\mathbf{z}_i}v_{\nu i}-\frac{dM^\mu(\mathbf{z}_i)}{ds_i}\right]=0.$$

Introducing the field function

$$A_i^\nu(\mathbf{x})=M^\nu(\mathbf{x})+\Sigma_{j\neq i}e_j\int_{s_j^0}^{s_j'}\delta(\mathbf{x}-\mathbf{z}_j)^2v_j^\nu ds_j, \tag{1·13}$$

it may be written

$$m_i\frac{dv_i^\mu}{ds_i}=e_i\left[\left(\frac{\partial A_i^\nu}{\partial x_\mu}\right)_{\mathbf{z}_i}v_{\nu i}-\frac{dA^\mu(\mathbf{z}_i)}{ds_i}\right]=e_i\left[\frac{\partial A_i^\nu}{\partial x_\mu}-\frac{\partial A_i^\mu}{\partial x_\nu}\right]_{\mathbf{z}_i}v_{\nu i}. \tag{1·14}$$

It is the correct Lorentz equation of motion of the ith particle, provided A_i^ν is connected with the ingoing and outgoing fields and the retarded and advanced fields of the other particles by the relation, given by the author (1938, equation (42)),

$$A_i^\nu=\tfrac{1}{2}[A_{\text{in}}^\nu+A_{\text{out}}^\nu]+\tfrac{1}{2}\Sigma_{j\neq i}[A_{j\,\text{ret}}^\nu+A_{j\,\text{adv}}^\nu]$$

or

$$A_i^\nu(\mathbf{x})=\tfrac{1}{2}[A_{\text{in}}^\nu(\mathbf{x})+A_{\text{out}}^\nu(\mathbf{x})]+\Sigma_{j\neq i}e_j\int_{-\infty}^{\infty}\delta(\mathbf{x}-\mathbf{z}_j)^2v_j^\nu ds_j. \tag{1·15}$$

According to (1·13) this requires

$$M^\nu(\mathbf{x})=\tfrac{1}{2}[A_{\text{in}}^\nu(\mathbf{x})+A_{\text{out}}^\nu(\mathbf{x})]+\Sigma_{j\neq i}e_j\left[\int_{-\infty}^{s_j^0}+\int_{s_j'}^{\infty}\right]\delta(\mathbf{x}-\mathbf{z}_j)^2v_j^\nu ds_j. \tag{1·16}$$

There is no need for equations (1·15), (1·16) to hold throughout space-time. It is sufficient to have them holding in the region which lies inside the future light-cone from $\mathbf{z}_i^0$ and inside the past light-cone from $\mathbf{z}_i'$, since it is only the value of $A_i^\nu(\mathbf{x})$ in this region that comes into play in the equation of motion (1·14) of the ith particle. In this region (1·16) may be replaced by

$$M^\nu(\mathbf{x})=\tfrac{1}{2}[A_{\text{in}}^\nu(\mathbf{x})+A_{\text{out}}^\nu(\mathbf{x})]+\Sigma_je_j\left[\int_{-\infty}^{s_j^0}+\int_{s'}^{\infty}\right]\delta(\mathbf{x}-\mathbf{z}_j)^2v_j^\nu ds_j, \tag{1·17}$$

with the sum taken over all values of j. By assuming (1·17) to hold throughout space-time one gets an expression for $M^\nu(\mathbf{x})$ independent of i, so that *the equations of motion of all the particles follow from the same action integral.* This is an important feature of Fokker's action principle.

Instead of (1·17) one may take

$$M^\nu(\mathbf{x}) = \tfrac{1}{2}[A^\nu_{\text{in}}(\mathbf{x}) + A^\nu_{\text{out}}(\mathbf{x})] + \tfrac{1}{2}\Sigma_j e_j \left[\int_{-\infty}^{s_j^0} - \int_{s_j'}^{\infty}\right] \Delta(\mathbf{x}-\mathbf{z}_j)\, v_j^\nu ds_j, \quad (1\cdot18)$$

where $\Delta(\mathbf{y})$ denotes the Jordan & Pauli (1928) Δ function of any 4-vector $\mathbf{y}$, connected with the δ function of $\mathbf{y}^2$ by

$$\Delta(\mathbf{y}) = \pm 2\delta(\mathbf{y}^2), \quad (1\cdot19)$$

the + or − sign being taken according to whether y_0 is positive or negative, respectively. (The Δ function used here and in equation (9) has the opposite sign to that of Jordan & Pauli.) The right-hand sides of (1·17) and (1·18) differ only inside the past light-cones from the points $\mathbf{z}_j^0$ and inside the future light-cones from the points $\mathbf{z}_j'$, and the field in these places does not enter into the equation of motion of any of the particles, on account of (1·6). (1·18) will be adopted here in preference to (1·17), since the former makes

$$\Box M^\nu(\mathbf{x}) = 0, \quad (1\cdot20)$$

on account of the general relation

$$\Box\Delta(\mathbf{x}) = 0 \quad (1\cdot21)$$

given by Jordan & Pauli (1928, equation (14)).

A difficulty appears in the work at this stage. In varying the action integral it was implicitly assumed, by the use of (1·11), that $M^\nu(\mathbf{x})$ is a continuous field function in the neighbourhood of the world-lines of the particles, but the expression subsequently obtained for it, (1·18), has discontinuities. One can easily see that those discontinuities on the world-line of any particle due to the terms in the sum in (1·18) referring to the other particles are not serious and do not invalidate the equations of motion (they produce the same sort of effects as a pulse in the incident radiation would). But those at the end-points of the world-line of any particle due to the term in the sum in (1·18) referring to that particle are serious, as they prevent the passage to the Hamiltonian formulation.

To get over the difficulty one must introduce a certain limiting process. Let $\boldsymbol{\lambda}$ be a small 4-vector whose direction is within the future light-cone, so that

$$\boldsymbol{\lambda}^2 > 0, \quad \lambda_0 > 0, \quad (1\cdot22)$$

and let the $\delta(\mathbf{z}_i - \mathbf{z}_j)^2$ in (1·2) be replaced by

$$\left.\begin{aligned} \delta^*(\mathbf{z}_i;\mathbf{z}_j) &= \delta(\mathbf{z}_i - \mathbf{z}_j + \boldsymbol{\lambda})^2 \quad \text{for} \quad (z_i - z_j)_0 > 0 \\ &= \delta(\mathbf{z}_i - \mathbf{z}_j - \boldsymbol{\lambda})^2 \quad \text{for} \quad (z_i - z_j)_0 < 0. \end{aligned}\right\} \quad (1\cdot23)$$

This will cause interaction to occur between pairs of particles at points on their world-lines that lie one just outside the light-cone from the other, instead of on it as previously. (It is assumed that two particles never approach to a distance of the order $\boldsymbol{\lambda}$ from one another.) One now has to replace δ by δ^* in all the equations as far as (1·17), and must replace (1·18) by

$$M^\nu(\mathbf{x}) = \tfrac{1}{2}[A^\nu_{\text{in}}(\mathbf{x})+A^\nu_{\text{out}}(\mathbf{x})] \\ +\tfrac{1}{2}\Sigma_j e_j\left[\int_{-\infty}^{s_j^0}\Delta(\mathbf{x}-\mathbf{z}_j+\boldsymbol{\lambda})\,v_j^\nu ds_j-\int_{s_j'}^{\infty}\Delta(\mathbf{x}-\mathbf{z}_j-\boldsymbol{\lambda})\,v_j^\nu ds_j\right], \quad (1·24)$$

the restrictions (1·6) having to be sharpened to

$$(\mathbf{z}_i^0-\mathbf{z}_j^0\pm\boldsymbol{\lambda})^2<0, \quad (\mathbf{z}_i'-\mathbf{z}_j'\pm\boldsymbol{\lambda})^2<0 \quad (i\neq j). \qquad (1·25)$$

The singularities in $M^\nu(\mathbf{x})$ are now displaced a little from the end-points of the world-lines and no longer come into play, even if the initial points $\mathbf{z}_j^0$ are also varied, provided the variations in the world-lines are smaller than $\boldsymbol{\lambda}$. The equations of motion resulting from the variation principle are no longer exact, but are correct only in the limit $\boldsymbol{\lambda}\to 0$. The limiting process here introduced appears to be unavoidable for setting up a satisfactory action theory and plays an important role in the later development, leading to the elimination of those divergent integrals (7) in quantum electrodynamics for which n is even.

One can now pass to the Hamiltonian formulation of the equations of motion. For each point in space-time $\mathbf{x}$, $M^\mu(\mathbf{x})$ may be counted as a co-ordinate, depending on the proper-times s_i' (it also depends on the proper-times s_i^0, but this does not here concern us), and will have a conjugate momentum, $K_\mu(\mathbf{x})$ say. These momenta, together with the particle momenta p_i^μ, are defined, as in the general theory of Weiss (1936), by the coefficients of $\partial M^\mu(\mathbf{x})$ and $Dz'_{\mu i}$ in the expression for ∂S, so that

$$\partial S = \Sigma_i p_i^\mu Dz'_{\mu i}+\iiiint_{-\infty}^{\infty} K_\mu(\mathbf{x})\,\partial M^\mu(\mathbf{x})\,dx_0dx_1dx_2dx_3. \qquad (1·26)$$

Comparing (1·26) with the sum of (1·9), (1·10) and (1·12), with δ^* written for δ, one obtains

$$K_\mu(\mathbf{x}) = \Sigma_i e_i\int_{s_i^0}^{s'}\delta(x_0-z_{0i})\,\delta(x_1-z_{1i})\,\delta(x_2-z_{2i})\,\delta(x_3-z_{3i})\,v_{\mu i}\,ds_i \quad (1·27)$$

and
$$p_i^\mu = m_i v_i'^\mu+e_i\left[M^\mu(\mathbf{z}_i')+\Sigma_{j\neq i}e_j\int_{s_j^0}^{s_j'}\delta^*(\mathbf{z}_i';\mathbf{z}_j)\,v_j^\mu ds_j\right]$$
$$= m_i v_i'^\mu+e_i\left[M^\mu(\mathbf{z}_i')+\tfrac{1}{2}\Sigma_j e_j\int_{s_j^0}^{s_j}\Delta(\mathbf{z}_i'-\mathbf{z}_j+\boldsymbol{\lambda})\,v_j^\mu ds_j\right], \qquad (1·28)$$

when one takes into account (1·23), (1·19) and the fact that a non-vanishing integrand can occur in (1·28) only when $j\neq i$ and $(\mathbf{z}_i'-\mathbf{z}_j+\boldsymbol{\lambda})_0>0$.

The momenta satisfy the Poisson Bracket relationships

$$[p_{\mu i}, z_{\nu j}] = g_{\mu\nu}\delta_{ij}, \tag{1·29}$$

$$[K_\mu(\mathbf{x}), M_\nu(\mathbf{x}')] = g_{\mu\nu}\delta(x_0-x'_0)\,\delta(x_1-x'_1)\,\delta(x_2-x'_2)\,\delta(x_3-x'_3), \tag{1·30}$$

and, of course, the Poisson Bracket of any two momenta or of any two co-ordinates vanishes. Instead of $K_\mu(\mathbf{x})$ it is more convenient to work with the momentum field function $N_\mu(\mathbf{x})$ defined by

$$N_\mu(\mathbf{x}) = \tfrac{1}{2}\iiiint_{-\infty}^{\infty} \Delta(\mathbf{x}-\mathbf{x}')\,K_\mu(\mathbf{x}')\,dx'_0dx'_1dx'_2dx'_3 \tag{1·31}$$

and satisfying

$$\Box N_\mu(\mathbf{x}) = 0 \tag{1·32}$$

in virtue of (1·21). Instead of (1·30) one has

$$[N_\mu(\mathbf{x}), M_\nu(\mathbf{x}')] = \tfrac{1}{2}g_{\mu\nu}\Delta(\mathbf{x}-\mathbf{x}'). \tag{1·33}$$

From (1·31) and (1·27)

$$N_\mu(\mathbf{x}) = \tfrac{1}{2}\Sigma_i e_i\int_{s_i^0}^{s'_i}\Delta(\mathbf{x}-\mathbf{z}_i)\,v_{\mu i}\,ds_i, \tag{1·34}$$

so that (1·28) may be written

$$\begin{aligned} p_i^\mu &= m_i v_i'^\mu + e_i[M^\mu(\mathbf{z}'_i)+N^\mu(\mathbf{z}'_i+\boldsymbol{\lambda})] \\ &= m_i v_i'^\mu + e_i A^\mu(\mathbf{z}'_i), \end{aligned} \tag{1·35}$$

where

$$A^\mu(\mathbf{x}) = M^\mu(\mathbf{x}) + N^\mu(\mathbf{x}+\boldsymbol{\lambda}). \tag{1·36}$$

From (1·20) and (1·32) the potentials $A_\mu(\mathbf{x})$ satisfy

$$\Box A_\mu(\mathbf{x}) = 0, \tag{1·37}$$

showing that they can be resolved into waves travelling with the velocity of light, and from (1·33)

$$[A_\mu(\mathbf{x}), A_\nu(\mathbf{x}')] = \tfrac{1}{2}g_{\mu\nu}\{\Delta(\mathbf{x}-\mathbf{x}'+\boldsymbol{\lambda})+\Delta(\mathbf{x}-\mathbf{x}'-\boldsymbol{\lambda})\}. \tag{1·38}$$

From (1·4) and (1·35)

$$F_i \equiv \{\mathbf{p}_i - e_i\mathbf{A}(\mathbf{z}'_i)\}^2 - m_i^2 = 0. \tag{1·39}$$

There is one of these equations for each particle. The expressions F_i may be used as Hamiltonians to determine how any dynamical variable ξ varies with the proper-times s'_i, in accordance with the equations

$$\kappa_i\,d\xi/ds'_i = [\xi, F_i]. \tag{1·40}$$

Here ξ is any function of the co-ordinates and momenta of the particles and of the fields **M**, **K**, **N**, **A**, and the κ's are multiplying factors not depending on ξ. Taking $\xi = z_i'^\mu$, one finds that

$$\kappa_i = -2m_i, \tag{1·41}$$

to get agreement with (1·35). Taking $\xi = p'^{\mu}_i$ gives one back the equation of motion (1·14) with the $\boldsymbol{\lambda}$ refinement. Taking $\xi = M_\mu(\mathbf{x})$, one gets

$$\begin{aligned} dM_\mu(\mathbf{x})/ds'_i &= e_i v'^{\nu}_i [M_\mu(\mathbf{x}), A_\nu(\mathbf{z}'_i)] \\ &= \tfrac{1}{2} e_i v'_{\mu i} \Delta(\mathbf{x} - \mathbf{z}'_i - \boldsymbol{\lambda}), \end{aligned} \qquad (1{\cdot}42)$$

from (1·36) and (1·33). This equation of motion for the field quantities $M_\mu(\mathbf{x})$ does not follow from the variation principle, as it involves only co-ordinates and velocities and not accelerations, and it has to be imposed as an extra condition in the variational method. With the method of development given here, it comes in as a consequence of the choice (1·24) for $M_\mu(\mathbf{x})$.

The above Hamiltonian formulation of the equations of classical electrodynamics may be taken over into the quantum theory in the usual way, by making the momenta into operators satisfying commutation relations corresponding to the Poisson Bracket relations (1·29), (1·30). Equation (1·38) in the limit $\boldsymbol{\lambda} \to 0$ goes over into the quantum equation (9). The Hamiltonians (1·39) provide the wave equations

$$F_i \psi = 0, \qquad (1{\cdot}43)$$

in which the wave function ψ is a function of the co-ordinates $\mathbf{z}'_i$ of all the particles, subject to the inequalities (1·25), and of the field variables $M_\mu(\mathbf{x})$. One can apply the theory to spinning electrons instead of spinless particles, by modifying the Hamiltonians F_i in the appropriate way.

The present form of quantum electrodynamics differs from that of Heisenberg & Pauli in two respects. First, it involves the $\boldsymbol{\lambda}$ limiting process. This limiting process may be introduced into the theory of Heisenberg & Pauli, as shown by the author (1939*b*), and leads to the elimination of the divergent integrals with even n-values. Secondly, it is expressed in terms of a different representation, with the wave functions involving the $M_\mu(\mathbf{x})$ as co-ordinates. This change in the representation eliminates the divergent integrals with odd n-values. The new representation cannot be interpreted in terms of only positive-energy photons, as is the case with the representation of Heisenberg & Pauli, but it can be interpreted in terms of both positive- and negative-energy photons, as will be shown in Appendix II.

APPENDIX II. RELATIVISTIC SECOND QUANTIZATION

The author's original method of second quantization for an assembly of similar Bose particles, consisting in the direct introduction as dynamical variables of the numbers of particles in the various quantum states for a particle, is not very suitable for the purpose of relativistic generalization. Fock (1934, §I) gave an improved method, which presents the basic ideas in a more direct manner and which allows itself to be easily developed into a relativistic theory. The main features of Fock's method will first be mentioned here, to provide a starting point. The author's (1939*a*) bracket notation will be used, as it seems to be the neatest and most concise one for this kind of work.

A state of an assembly of u particles satisfying the Bose statistics is represented by a symmetrical wave function $\langle q'\,q''\dots q^u|\rangle$, where each value for a variable q represents a state for one particle. Let us suppose for definiteness that the values a q can take on are discrete. Then $|\langle q'\,q''\dots q^u|\rangle|^2$ is the probability of the first particle being in the state q', the second in the state q'', and so on. Since, however, there is no means of distinguishing one particle from another, the only physically significant interpretation of the wave function is that, if n^a, n^b, ... are the numbers of particles in the various states $q^a, q^b, \dots$ when the distribution $q', q'', \dots, q^u$ occurs, so that $\Sigma n = u$, then

$$(u!/n^a!\,n^b!\dots)\,|\langle q'\,q''\dots q^u|\rangle|^2 \tag{2·1}$$

is the total probability of there being these numbers of particles in the various states.

If the number of particles in the assembly is indeterminate, a state of the assembly must be represented by a series of component wave functions

$$\langle\cdot|\rangle,\ \langle q'|\rangle,\ \langle q'\,q''|\rangle,\ \langle q'\,q''\,q'''|\rangle,\ \dots \tag{2·2}$$

referring successively to $0, 1, 2, 3, \dots$ particles, those referring to two or more particles being symmetrical. Each component $\langle q'\,q''\dots q^u|\rangle$ determines the probability of there being u particles distributed in any way over the various states, in accordance with (2·1). $\Sigma_{q'\dots q^u}|\langle q'\,q''\dots q^u|\rangle|^2$ is the total probability of there being u particles in the assembly, and this quantity summed for all u must equal unity.

Fock's method consists in introducing a set of variables $\xi_{q'}, \xi_{q''}, \dots$, one for each value of q or for each independent state of a particle, and representing a state of the assembly by the function $\langle\xi|\rangle$ of the ξ's

$$\langle\xi|\rangle = \Sigma_u\, u!^{-\frac{1}{2}}\, \Sigma_{q'\dots q^u}\, \xi_{q'}\, \xi_{q''} \dots \xi_{q^u} \langle q'\,q''\dots q^u|\rangle. \tag{2·3}$$

This function is a polynomial in the ξ's, the terms of any degree u corresponding to there being u particles in the assembly. The terms of the form $\xi_{q^a}^{n^a}\,\xi_{q^b}^{n^b}\dots$ correspond to there being n^a particles in the state q^a, n^b in the state q^b, and so on, and hence the operator $\xi_{q^a}\,\partial/\partial\xi_{q^a}$ gives the number of particles n^a in any state q^a. From (2·1) it is easily seen that the probability of there being n^a particles in the state q^a, n^b in the state q^b, and so on, is

$$n^a!\,n^b!\ldots\left|\text{coefficient of } \xi_{q^a}^{n^a}\xi_{q^b}^{n^b}\ldots \text{ in } \langle\xi|\rangle\right|^2. \tag{2·4}$$

The transformation from the set of components (2·2) to the function $\langle\xi|\rangle$ is a linear transformation and is therefore expressible in terms of a transformation function $\langle\xi|q'q''\ldots q^u\rangle$, thus

$$\langle\xi|\rangle = \Sigma_u\,\Sigma_{q'\ldots q^u}\langle\xi|q'\,q''\ldots q^u\rangle\,\langle q'\,q''\ldots q^u|\rangle. \tag{2·5}$$

Comparing (2·5) with (2·3), one finds

$$\langle\xi|q'\,q''\ldots q^u\rangle = u!^{-\frac{1}{2}}\,\xi_{q'}\,\xi_{q''}\ldots\xi_{q^u}. \tag{2·6}$$

The operators ξ_q, $\partial/\partial\xi_q$, when operating to the right, correspond to the emission of a particle into and the absorption of a particle from the state q, respectively. Let us see what these operators go over into when transformed to the q representation (2·2). Using (2·6) and the laws of the transformation theory, one gets

$$\begin{aligned}\langle\xi|\xi_q|q'\ldots q^u\rangle = \xi_q\langle\xi|q'\ldots q^u\rangle &= u!^{-\frac{1}{2}}\,\xi_q\,\xi_{q'}\ldots\xi_{q^u}\\ &= (u+1)^{\frac{1}{2}}\,\langle\xi|q\,q'\ldots q^u\rangle,\end{aligned}$$

so that

$$\xi_q|q'\ldots q^u\rangle = (u+1)^{\frac{1}{2}}\,|q\,q'\ldots q^u\rangle. \tag{2·7}$$

This shows that the emission operator ξ_q, operating on one of the basic $\rangle$ vectors of the q representation referring to a state of u particles, changes it into a basic $\rangle$ vector referring to a state of $(u+1)$ particles, with the coefficient $(u+1)^{\frac{1}{2}}$. Again

$$\begin{aligned}\langle\xi|\frac{\partial}{\partial\xi_q}|q'\ldots q^u\rangle = \frac{\partial}{\partial\xi_q}\langle\xi|q'\ldots q^u\rangle &= u!^{-\frac{1}{2}}\frac{\partial}{\partial\xi_q}\xi_{q'}\ldots\xi_{q^u}\\ &= u!^{-\frac{1}{2}}\,\Sigma_{r=1}^{u}\,\delta_{q\,q^r}\,\xi_{q'}\ldots\xi_{q^{r-1}}\,\xi_{q^{r+1}}\ldots\xi_{q^u}\\ &= u^{-\frac{1}{2}}\,\Sigma_r\,\delta_{q\,q^r}\langle\xi|q'\ldots q^{r-1}\,q^{r+1}\ldots q^u\rangle,\end{aligned}$$

so that

$$\frac{d}{d\xi_q}|q'\ldots q^u\rangle = u^{-\frac{1}{2}}\,\Sigma_r\,\delta_{q\,q^r}|q'\ldots q^{r-1}\,q^{r+1}\ldots q^u\rangle. \tag{2·8}$$

This shows the effect of the absorption operator $d/d\xi_q$ operating on one of the basic $\rangle$ vectors of the q representation.

One can now verify that ξ_q and $\partial/\partial\xi_q$ are adjoint operators when expressed in the q representation. One must make use of

$$\langle q'\,q''\ldots q^u|q^{u+1}\,q^{u+2}\ldots q^{2u}\rangle = u!^{-1}\,\delta(q'\,q''\ldots q^u;\,q^{u+1}\,q^{u+2}\ldots q^{2u}), \tag{2·9}$$

where $\delta(q'\,q''\ldots q^u;\,q^{u+1}\,q^{u+2}\ldots q^{2u})$ equals unity if q^{u+1}, $q^{u+2}\ldots q^{2u}$ are any permutation of $q', q'', \ldots, q^u$ and zero otherwise. The coefficient $u!^{-1}$ on the right of (2·9) is needed to make $\langle q' q'' \ldots q^u | q^{u+1} q^{u+2} \ldots q^{2u} \rangle$ equivalent to the unit matrix when multiplied into symmetrical functions of u variables q by matrix multiplication. From (2·7)

$$\begin{aligned}\langle q^{u+1} q^{u+2} \ldots q^{2u+1} | \xi_q | q' \ldots q^u \rangle & \\ &= (u+1)^{\frac{1}{2}} \langle q^{u+1} q^{u+2} \ldots q^{2u+1} | q\, q' \ldots q^u \rangle \\ &= (u+1)^{\frac{1}{2}} (u+1)!^{-1} \delta(q^{u+1} q^{u+2} \ldots q^{2u+1};\, q\, q' \ldots q^u), \qquad (2\cdot10)\end{aligned}$$

and from (2·8)

$$\begin{aligned}\langle q' \ldots q^u | \partial/\partial\xi_q | q^{u+1} \ldots q^{2u+1} \rangle & \\ &= (u+1)^{-\frac{1}{2}} \textstyle\sum_{r=u+1}^{2u+1} \delta_{q\,q^r} \langle q' \ldots q^u | q^{u+1} \ldots q^{r-1} q^{r+1} \ldots q^{2u+1} \rangle \\ &= (u+1)^{-\frac{1}{2}}\, u!^{-1} \textstyle\sum_{r=u+1}^{2u+1} \delta_{qq^r}\, \delta(q' \ldots q^u;\, q^{u+1} \ldots q^{r-1} q^{r+1} \ldots q^{2u+1}) \\ &= (u+1)^{-\frac{1}{2}}\, u!^{-1} \delta(q\, q' \ldots q^u;\, q^{u+1} \ldots q^{2u+1}). \qquad (2\cdot11)\end{aligned}$$

The expressions (2·10) and (2·11) are real and equal and the other kinds of matrix elements of ξ_q and $\partial/\partial\xi_q$ all vanish, and hence ξ_q and $\partial/\partial\xi_q$ are adjoint.

The above theory may be made relativistic in a straightforward way. For particles with no spin, one replaces each q by $\mathbf{x}$, denoting the four coordinates of the particle in space-time. Thus one gets, instead of (2·2), the series of component wave functions

$$\langle\cdot|\rangle,\ \langle \mathbf{x}'|\rangle,\ \langle \mathbf{x}'\mathbf{x}''|\rangle,\ \langle \mathbf{x}'\mathbf{x}''\mathbf{x}'''|\rangle,\ \ldots, \qquad (2\cdot12)$$

those beyond the second being symmetrical. Let us take for simplicity the case of particles of zero rest-mass. Then the various components (2·12) satisfy

$$\Box_{\mathbf{x}^r} \langle \mathbf{x}'\mathbf{x}'' \ldots \mathbf{x}^u|\rangle = 0, \quad \mathbf{x}^r = \mathbf{x}', \mathbf{x}'', \ldots, \mathbf{x}^u. \qquad (2\cdot13)$$

One now has to incorporate the Gordon-Klein rule (5) for the probability. This requires one to take for the scalar product $\langle a|b\rangle$ of a wave function $\langle \mathbf{x}|b\rangle$ and a conjugate complex wave function $\langle a|\mathbf{x}\rangle$ referring to a single particle

$$\langle a|b\rangle = \frac{1}{4\pi i} \iiint \left[\frac{\partial}{\partial x_0} \langle a|\mathbf{x}\rangle \,.\, \langle \mathbf{x}|b\rangle - \langle a|\mathbf{x}\rangle \frac{\partial}{\partial x_0} \langle \mathbf{x}|b\rangle\right] dx_1 dx_2 dx_3,$$

the integral being taken over any three-dimensional surface x_0 = constant. This result may be written for brevity

$$\langle a|b\rangle = (4\pi i)^{-1} \int \langle a|\mathbf{x}\rangle (\reflectbox{Γ} - \Gamma) \langle \mathbf{x}|b\rangle\, dx, \qquad (2\cdot14)$$

where dx means $dx_1 dx_2 dx_3$ and $\reflectbox{$\Gamma$}$ and Γ mean $\partial/\partial x_0$ operating on the expression to the left and to the right respectively.

As examples of (2·14), it follows that

$$\langle a|\mathbf{x}''\rangle = (4\pi i)^{-1}\int \langle a|\mathbf{x}'\rangle\,(\text{⅂}'-\Gamma')\,\langle \mathbf{x}'|\mathbf{x}''\rangle\,dx'$$

and

$$\langle \mathbf{x}'|b\rangle = (4\pi i)^{-1}\int \langle \mathbf{x}'|\mathbf{x}''\rangle\,(\text{⅂}''-\Gamma'')\,\langle \mathbf{x}''|b\rangle\,dx''.$$

These results, considered as equations for the unknown function $\langle \mathbf{x}'|\mathbf{x}''\rangle$, have the solution

$$\langle \mathbf{x}'|\mathbf{x}''\rangle = -i\Delta(\mathbf{x}'-\mathbf{x}''), \tag{2·15}$$

as is easily seen by taking the integrals in them to be over the three-dimensional surface $x_0' = x_0''$ and using the property of the Δ function given by Heisenberg & Pauli (1929*a*, equation (64))

$$\frac{\partial\Delta(\mathbf{x})}{\partial x_0} = 4\pi\,\delta(x_1)\,\delta(x_2)\,\delta(x_3) \quad \text{for} \quad x_0 = 0. \tag{2·16}$$

Equation (2·15), together with

$$\Box\langle \mathbf{x}| = 0, \quad \Box|\mathbf{x}\rangle = 0, \tag{2·17}$$

give the relations between the basic vectors $|\mathbf{x}\rangle$, $\langle \mathbf{x}|$ of the present representation for one particle, replacing the conditions of the non-relativistic theory that these vectors are all normalized and orthogonal to one another.

If there are u particles present, equation (2·14) must be generalized to

$$\langle a|b\rangle = (4\pi i)^{-u}\iint_{\cdots} \langle a|\mathbf{x}'\mathbf{x}''\ldots\mathbf{x}^u\rangle\,(\text{⅂}'-\Gamma')\,(\text{⅂}''-\Gamma'')\ldots(\text{⅂}^u-\Gamma^u)\,\langle \mathbf{x}'\mathbf{x}''\ldots\mathbf{x}^u|b\rangle\,dx'\,dx''\ldots dx^u, \tag{2·18}$$

in which the various ⅂ and Γ operators all commute with one another. If the number of particles is indeterminate, the right-hand side of (2·18) must be summed for all values of u. If the states a and b are the same and the wave functions are properly normalized, the integrand in (2·18), including the coefficient $(4\pi i)^{-u}$, gives the probability of there being u particles in the neighbourhoods of the points $\mathbf{x}', \mathbf{x}'', \ldots, \mathbf{x}^u$, per unit volume of each neighbourhood. Corresponding to (2·9), equation (2·15) must be generalized to

$$\langle \mathbf{x}'\mathbf{x}''\ldots\mathbf{x}^u|\mathbf{x}^{u+1}\mathbf{x}^{u+2}\ldots\mathbf{x}^{2u}\rangle = (-i)^u\,u!^{-1}\,\Sigma\Delta(\mathbf{x}'-\mathbf{x}^a)\,\Delta(\mathbf{x}''-\mathbf{x}^b)\ldots\Delta(\mathbf{x}^u-\mathbf{x}^z), \tag{2·19}$$

in which $a, b, \ldots, z$ denotes some permutation of $u+1, u+2, \ldots, 2u$ and the summation is over all such permutations.

One can proceed to set up the relativistic analogue of the fundamental transformation (2·3). Introducing a set of variables $\xi_{\mathbf{x}}$, one for each point in space-time, and assuming the transformation function

$$\langle \xi|\mathbf{x}'\mathbf{x}''\ldots\mathbf{x}^u\rangle = u!^{-\frac{1}{2}}\,\xi_{\mathbf{x}'}\,\xi_{\mathbf{x}''}\ldots\xi_{\mathbf{x}^u} \tag{2·20}$$

as the analogue of (2·6), one gets

$$\langle\xi|\rangle = \Sigma_u (4\pi i)^{-u} \iint_{\dots} \langle\xi|\mathbf{x}'\mathbf{x}''\dots\mathbf{x}^u\rangle (\mathrm{T}'-\Gamma')\dots(\mathrm{T}^u-\Gamma^u)\langle\mathbf{x}'\dots\mathbf{x}^u|\rangle\, dx'\dots dx^u$$

$$= \Sigma_u (4\pi i)^{-u}\, u!^{-\frac{1}{2}} \iint_{\dots} \xi_{\mathbf{x}'}\dots\xi_{\mathbf{x}^u}(\mathrm{T}'-\Gamma')\dots(\mathrm{T}^u-\Gamma^u)\langle\mathbf{x}'\dots\mathbf{x}^u|\rangle\, dx'\dots dx^u. \tag{2·21}$$

The analogue of (2·7) is

$$\xi_{\mathbf{x}}|\mathbf{x}'\dots\mathbf{x}^u\rangle = (u+1)^{\frac{1}{2}}\,|\mathbf{x}\,\mathbf{x}'\dots\mathbf{x}^u\rangle, \tag{2·22}$$

showing that $\xi_{\mathbf{x}}$ corresponds to the emission of a particle at the point $\mathbf{x}$ in space-time. Since $\Box|\mathbf{x}\,\mathbf{x}'\dots\mathbf{x}^u\rangle = 0$, as follows from the conjugate imaginary equation to (2·13) applied to $(u+1)$ particles, one has

$$\Box\xi_{\mathbf{x}} = 0. \tag{2·23}$$

This shows that the $\xi_{\mathbf{x}}$ are not all independent, as is to be expected physically.

The lack of independence of the $\xi_{\mathbf{x}}$ prevents one from giving a direct meaning to $\partial/\partial\xi_{\mathbf{x}}$. One can, however, get an operator to play the part of the absorption operator by taking the adjoint $\xi^*_{\mathbf{x}}$ to $\xi_{\mathbf{x}}$ in the $\mathbf{x}$ representation. With the help of (2·22) and (2·19)

$$\langle\mathbf{x}'\dots\mathbf{x}^u|\xi_{\mathbf{x}}|\mathbf{x}^{u+1}\dots\mathbf{x}^{2u-1}\rangle = u^{\frac{1}{2}}\langle\mathbf{x}'\dots\mathbf{x}^u|\mathbf{x}\,\mathbf{x}^{u+1}\dots\mathbf{x}^{2u-1}\rangle$$

$$= (-i)^u u^{\frac{1}{2}} u!^{-1}\,\Sigma\Delta(\mathbf{x}^a-\mathbf{x})\,\Delta(\mathbf{x}^b-\mathbf{x}^{u+1})\dots\Delta(\mathbf{x}^z-\mathbf{x}^{2u-1}), \tag{2·24}$$

where $\mathbf{x}^a, \mathbf{x}^b, \dots, \mathbf{x}^z$ denotes some permutation of $\mathbf{x}', \mathbf{x}'', \dots, \mathbf{x}^u$ and the summation is over all such permutations. The sum in (2·24) may be replaced by

$$\Sigma_{r=1}^{u}\,\Delta(\mathbf{x}^r-\mathbf{x})\,\Sigma\Delta(\mathbf{x}^a-\mathbf{x}^{u+1})\dots\Delta(\mathbf{x}^z-\mathbf{x}^{2u-1}), \tag{2·25}$$

where $\mathbf{x}^a\dots\mathbf{x}^z$ now denotes some permutation of $\mathbf{x}'\dots\mathbf{x}^{r-1}\mathbf{x}^{r+1}\dots\mathbf{x}^u$ and the second Σ refers to all such permutations. Expression (2·25) equals

$$\Sigma_r\Delta(\mathbf{x}^r-\mathbf{x})\,i^{u-1}(u-1)!\,\langle\mathbf{x}'\dots\mathbf{x}^{r-1}\mathbf{x}^{r+1}\dots\mathbf{x}^u|\mathbf{x}^{u+1}\dots\mathbf{x}^{2u-1}\rangle$$

from (2·19). Substituting this for the sum in (2·24) and taking the conjugate complex of both sides of the equation, one obtains

$$\langle\mathbf{x}^{u+1}\dots\mathbf{x}^{2u-1}|\xi^*_{\mathbf{x}}|\mathbf{x}'\dots\mathbf{x}^u\rangle$$

$$= iu^{-\frac{1}{2}}\,\Sigma_r\Delta(\mathbf{x}^r-\mathbf{x})\,\langle\mathbf{x}^{u+1}\dots\mathbf{x}^{2u-1}|\mathbf{x}'\dots\mathbf{x}^{r-1}\mathbf{x}^{r+1}\dots\mathbf{x}^u\rangle$$

and hence

$$\xi^*_{\mathbf{x}}|\mathbf{x}'\dots\mathbf{x}^u\rangle = -iu^{-\frac{1}{2}}\,\Sigma_r\Delta(\mathbf{x}-\mathbf{x}^r)|\mathbf{x}'\dots\mathbf{x}^{r-1}\mathbf{x}^{r+1}\dots\mathbf{x}^u\rangle. \tag{2·26}$$

This is the relativistic analogue of (2·8). It shows, from (1·21), that

$$\Box\xi^*_{\mathbf{x}} = 0. \tag{2·27}$$

Let us determine the commutation relation connecting $\xi_{\mathbf{x}}$ and $\xi_{\mathbf{x}}^*$. From (2·22) and (2·26)

$$\begin{aligned}\xi_{\mathbf{x}}^*\,\xi_{\mathbf{x}^0}|\mathbf{x}'\ldots\mathbf{x}^u\rangle &= (u+1)^{\frac{1}{2}}\,\xi_{\mathbf{x}}^*|\mathbf{x}^0\mathbf{x}'\ldots\mathbf{x}^u\rangle\\ &= -i\Sigma_{r=0}^{u}\,\Delta(\mathbf{x}-\mathbf{x}^r)|\mathbf{x}^0\mathbf{x}'\ldots\mathbf{x}^{r-1}\mathbf{x}^{r+1}\ldots\mathbf{x}^u\rangle,\end{aligned}$$

and again

$$\begin{aligned}\xi_{\mathbf{x}^0}\,\xi_{\mathbf{x}}^*|\mathbf{x}'\ldots\mathbf{x}^u\rangle &= -iu^{-\frac{1}{2}}\,\xi_{\mathbf{x}^0}\,\Sigma_{r=1}^{u}\,\Delta(\mathbf{x}-\mathbf{x}^r)|\mathbf{x}'\ldots\mathbf{x}^{r-1}\mathbf{x}^{r+1}\ldots\mathbf{x}^u\rangle\\ &= -i\Sigma_{r=1}^{u}\,\Delta(\mathbf{x}-\mathbf{x}^r)|\mathbf{x}^0\mathbf{x}\ldots\mathbf{x}'^{r-1}\mathbf{x}^{r+1}\ldots\mathbf{x}^u\rangle.\end{aligned}$$

Hence $$(\xi_{\mathbf{x}}^*\,\xi_{\mathbf{x}^0}-\xi_{\mathbf{x}^0}\,\xi_{\mathbf{x}}^*)|\mathbf{x}'\ldots\mathbf{x}^u\rangle = -i\,\Delta(\mathbf{x}-\mathbf{x}^0)|\mathbf{x}'\ldots\mathbf{x}^u\rangle,$$

so that $$\xi_{\mathbf{x}}^*\,\xi_{\mathbf{x}^0}-\xi_{\mathbf{x}^0}\,\xi_{\mathbf{x}}^* = -i\,\Delta(\mathbf{x}-\mathbf{x}^0). \qquad (2·28)$$

This result may be regarded as the analogue in second quantization of the result (2·15) for a single particle.

The above relativistic theory for spinless particles of zero rest-mass may easily be extended to apply to photons. The wave function for a single particle must be made into a 4-vector with components corresponding to the four components of the electromagnetic potential. The 4-vector wave function may be considered as a function of the four coordinates $\mathbf{x}$ and of another variable μ taking on the four values 0, 1, 2, 3, and may be written $\langle\mathbf{x}_\mu|\rangle$. The μ here is not a suffix attached to $\mathbf{x}$, but is an independent variable, which is written in the lower position to express that the whole function $\langle\mathbf{x}_\mu|\rangle$ is subject to the contravariant law under Lorentz transformations, and which may be raised by the usual rule

$$\langle\mathbf{x}^\mu|\rangle = g^{\mu\nu}\langle\mathbf{x}_\nu|\rangle. \qquad (2·29)$$

The multiplication rule (2·14) must be extended to

$$\langle a|b\rangle = -g^{\mu\nu}(4\pi i)^{-1}\int\langle a|\mathbf{x}_\mu\rangle\,(\text{⅂}-\Gamma)\,\langle\mathbf{x}_\nu|b\rangle\,dx, \qquad (2·30)$$

in which the $-$ sign is inserted with $g^{\mu\nu}$ for convenience, so as to give a factor $+1$ for μ and ν equal to 1, 2 or 3. Similarly, (2·15) must be extended to

$$\langle\mathbf{x}'_\mu|\mathbf{x}''_\nu\rangle = ig_{\mu\nu}\,\Delta(\mathbf{x}'-\mathbf{x}''). \qquad (2·31)$$

In a corresponding way the wave function for u particles must be made into a tensor of rank u and may be written $\langle\mathbf{x}'_\mu\mathbf{x}''_\nu\ldots\mathbf{x}^u_\tau|\rangle$, in which the $\mu, \nu, \ldots, \tau$ are independent variables from the $\mathbf{x}$'s. It must satisfy the symmetry condition that it remains unchanged when any permutation is applied to the $\mathbf{x}$'s and the same permutation to $\mu, \nu, \ldots, \tau$. The multiplication rule (2·18) must be extended to

$$\langle a|b\rangle = (-4\pi i)^{-u} \iint_{\cdots} \langle a|\mathbf{x}'_\mu \mathbf{x}''_\nu \ldots \mathbf{x}^u_\tau\rangle (\Gamma' - \Gamma')(\Gamma'' - \Gamma'') \ldots (\Gamma^u - \Gamma^u) \langle \mathbf{x}'^\mu \mathbf{x}''^\nu \ldots \mathbf{x}^{u\tau}|b\rangle \, dx' \, dx'' \ldots dx^u. \quad (2{\cdot}32)$$

The further development of these extensions is obvious and need not be mentioned in detail. The result is that the emission operator becomes a 4-vector $\xi_{\mathbf{x}\mu}$, and satisfies with its adjoint $\xi^*_{\mathbf{x}\mu}$ in the $\mathbf{x}_\mu$ representation the commutation relation

$$\xi^*_{\mathbf{x}\mu} \xi_{\mathbf{x}'\nu} - \xi_{\mathbf{x}'\nu} \xi^*_{\mathbf{x}\mu} = i g_{\mu\nu} \Delta(\mathbf{x} - \mathbf{x}'), \quad (2{\cdot}33)$$

corresponding to (2·28).

To establish a connexion between the present theory of photons and the quantum electrodynamics of the end of Appendix I, one must put

$$M_\mu(\mathbf{x}) = (\tfrac{1}{2}\hbar)^{\frac{1}{2}} \xi_{\mathbf{x}\mu}, \quad N_\mu(\mathbf{x}) = (\tfrac{1}{2}\hbar)^{\frac{1}{2}} \xi^*_{\mathbf{x}\mu}, \quad (2{\cdot}34)$$

so as to make the commutation relation (1·33) go over into (2·33). The wave function ψ of Appendix I, which is of the form of a power series in the variables $M_\mu(\mathbf{x})$, may now be identified with $\langle \xi | \rangle$. The interpretation of ψ that one would expect from the rules of non-relativistic quantum mechanics, namely, that $|\psi|^2$ gives the probability of the variables $M_\mu(\mathbf{x})$ having specified values, is no longer applicable. Instead one must work from (2·4) suitably modified to fit in with the Gordon-Klein expression for the probability for one particle. One may make a Fourier resolution of the $\xi_{\mathbf{x}\mu}$ and replace the continuous range of Fourier components by a discrete set. The wave function then becomes a power series in ξ variables referring to the Fourier components and its coefficients will give, according to the formula (3·24) of Appendix III, the probabilities of there being various numbers of photons in various momentum states, some with positive and some with negative energy. These probabilities are then to be re-interpreted, in accordance with one of the fundamental assumptions of the text, as the probabilities of various photons having been emitted or absorbed.

The real dynamical variables $M_\mu(\mathbf{x})$ appear in the present theory as operators $(\frac{1}{2}\hbar)^{\frac{1}{2}} \xi_{\mathbf{x}\mu}$ which are not self-adjoint. This marks a departure from one of the usual rules of quantum mechanics. However, the Hamiltonians F_t are self-adjoint in the limit $\lambda \to 0$, since from (1·39) and (1·36) they involve the field variables only through the linear combination $M_\mu(\mathbf{x}) + N_\mu(\mathbf{x})$ or $\xi_{\mathbf{x}\mu} + \xi^*_{\mathbf{x}\mu}$. This result is sufficient to ensure the conservation law—that if the wave function is initially normalized it always remains so—and makes the existence of other real dynamical variables not corresponding to self-adjoint operators unimportant.

APPENDIX III. THE TRANSFORMATION CONNECTING THE TWO FORMS OF QUANTUM ELECTRODYNAMICS

The new form of quantum electrodynamics set up in Appendices I and II and the old form of Heisenberg & Pauli (amended by the introduction of the $\boldsymbol{\lambda}$ limiting process) involve two different representations of the same set of operators describing the electromagnetic field, and one can discuss the transformation from one of these representations to the other on the lines of the general transformation theory of quantum mechanics. One cannot, as before mentioned, carry out the transformation for solutions of the wave equations, since such solutions in the old representation contain divergent integrals, but one can carry it out for other representatives of states of the field of a sufficiently general kind to make the methods of the transformation theory applicable.

It is convenient to work with the Fourier components of the various field quantities. With the new form of quantum electrodynamics, the emission operator $\xi_{\mathbf{x}\mu}$ satisfies (2·23) and its Fourier resolution is like that of the wave function representing a single photon and reads

$$\xi_{\mathbf{x}\mu} = (2\pi)^{-1}\Sigma\iiint \xi_{\mathbf{k}\mu} e^{i(\mathbf{k},\mathbf{x})} k_0^{-1} dk_1 dk_2 dk_3, \tag{3·1}$$

where the Σ means a summation over both values $\pm\sqrt{(k_1^2+k_2^2+k_3^2)}$ for k_0, as in equation (12). The Fourier coefficient $\xi_{\mathbf{k}\mu}$ introduced by (3·1) is defined only for $\mathbf{k}^2 = 0$, and is the operator of emission of a photon with energy and momentum given by the 4-vector $\mathbf{k}\hbar$. Similarly the Fourier resolution of the absorption operator $\xi^*_{\mathbf{x}\mu}$ reads

$$\xi^*_{\mathbf{x}\mu} = (2\pi)^{-1}\Sigma\iiint \xi^*_{\mathbf{k}\mu} e^{-i(\mathbf{k},\mathbf{x})} k_0^{-1} dk_1 dk_2 dk_3, \tag{3·2}$$

where $\xi^*_{\mathbf{k}\mu}$, the adjoint of $\xi_{\mathbf{k}\mu}$, is the operator of absorption of a photon with energy and momentum $\mathbf{k}\hbar$. The commutation relation (2·33) connecting $\xi_{\mathbf{x}\mu}$ and $\xi^*_{\mathbf{x}\mu}$ leads to the following commutation relation connecting $\xi_{\mathbf{k}\mu}$ and $\xi^*_{\mathbf{k}\mu}$

$$\left.\begin{aligned} \xi^*_{\mathbf{k}\mu}\xi_{\mathbf{k}'\nu} - \xi_{\mathbf{k}'\nu}\xi^*_{\mathbf{k}\mu} &= -g_{\mu\nu}\,k_0\,\delta(k_1-k_1')\,\delta(k_2-k_2')\,\delta(k_3-k_3') \\ &\qquad\qquad \text{when } k_0,\ k_0' \text{ have the same sign,} \\ &= 0 \qquad\quad \text{when } k_0,\ k_0' \text{ have opposite signs.} \end{aligned}\right\} \tag{3·3}$$

The potentials A_μ occurring in the Hamiltonians are, according to (1·36) and (2·34), in the limit $\boldsymbol{\lambda}\to 0$,

$$A_\mu(\mathbf{x}) = M_\mu(\mathbf{x}) + N_\mu(\mathbf{x}) = (\tfrac{1}{2}\hbar)^{\frac{1}{2}}(\xi_{\mathbf{x}\mu} + \xi^*_{\mathbf{x}\mu}),$$

and thus have the Fourier resolution

$$A_\mu(\mathbf{x}) = (\tfrac{1}{2}\hbar)^{\frac{1}{2}}\,(2\pi)^{-1}\,\Sigma\iiint\{\xi_{\mathbf{k}\mu}\,e^{i(\mathbf{k},\mathbf{x})}+\xi^*_{\mathbf{k}\mu}\,e^{-i(\mathbf{k},\mathbf{x})}\}\,k_0^{-1}\,dk_1dk_2dk_3. \tag{3·4}$$

This agrees with $\mathbf{A}^3(\mathbf{x})$ of equation (12), provided one takes

$$R_{\mathbf{k}\mu} = \hbar^{\frac{1}{2}}(2\pi)^{-1}\,\xi_{\mathbf{k}\mu}, \quad \bar{R}_{\mathbf{k}\mu} = \hbar^{\frac{1}{2}}(2\pi)^{-1}\,\xi^*_{\mathbf{k}\mu}. \tag{3·5}$$

The potentials B_μ defined by

$$B_\mu(\mathbf{x}) = M_\mu(\mathbf{x}) - N_\mu(\mathbf{x}) = (\tfrac{1}{2}\hbar)^{\frac{1}{2}}\,(\xi_{\mathbf{x}\mu}-\xi^*_{\mathbf{x}\mu})$$

commute with the A's, as is easily seen from (2·33), and are thus constants of the motion and also redundant variables. They have the Fourier resolution

$$B_\mu(\mathbf{x}) = (\tfrac{1}{2}\hbar)^{\frac{1}{2}}\,(2\pi)^{-1}\,\Sigma\iiint\{\xi_{k\mu}\,e^{i(\mathbf{k},\mathbf{x})}-\xi^*_{\mathbf{k}\mu}\,e^{-i(\mathbf{k},\mathbf{x})}\}\,k_0^{-1}\,dk_1dk_2dk_3. \tag{3·6}$$

From (3·5) one sees that the present B_μ is the same as that of equation (14). The $B_\mu(\mathbf{x})$ satisfy a similar commutation relation to the $A_\mu(\mathbf{x})$, equation (9), with the opposite sign on the right.

With the old form of quantum electrodynamics of Heisenberg & Pauli, the Fourier resolution of the potentials A_μ occurring in the Hamiltonians must be effected in the form

$$\left.\begin{aligned} A_\mu(\mathbf{x}) &= \hbar^{\frac{1}{2}}(2\pi)^{-1}\iiint\{\alpha_{\mathbf{k}\mu}\,e^{i(\mathbf{k},\mathbf{x})}+\bar{\alpha}_{\mathbf{k}\mu}\,e^{-i(\mathbf{k},\mathbf{x})}\}\,k_0^{-1}\,dk_1dk_2dk_3,\\ k_0 &= \sqrt{(k_1^2+k_2^2+k_3^2)},\end{aligned}\right\} \tag{3·7}$$

in which $\alpha_{\mathbf{k}\mu}$, $\bar{\alpha}_{\mathbf{k}\mu}$, defined only for positive k_0, are considered as the operators of emission into and absorption from positive-energy photon states, and satisfy the commutation relation

$$\bar{\alpha}_{\mathbf{k}\mu}\alpha_{\mathbf{k}'\nu}-\alpha_{\mathbf{k}'\nu}\bar{\alpha}_{\mathbf{k}\mu} = -g_{\mu\nu}\,k_0\,\delta(k_1-k_1')\,\delta(k_2-k_2')\,\delta(k_3-k_3'). \tag{3·8}$$

The redundant variables $B_\mu(\mathbf{x})$, if introduced into this theory, may correspondingly be resolved into Fourier components in the form

$$\left.\begin{aligned} B_\mu(\mathbf{x}) &= \hbar^{\frac{1}{2}}(2\pi)^{-1}\iiint\{\bar{\beta}_{\mathbf{k}\mu}\,e^{i(\mathbf{k},\mathbf{x})}+\beta_{\mathbf{k}\mu}\,e^{-i(\mathbf{k},\mathbf{x})}\}\,k_0^{-1}\,dk_1dk_2dk_3,\\ k_0 &= \sqrt{(k_1^2+k_2^2+k_3^2)},\end{aligned}\right\} \tag{3·9}$$

with $\beta_{\mathbf{k}\mu}$, $\bar{\beta}_{\mathbf{k}\mu}$, which are defined only for positive k_0, commuting with $\alpha_{\mathbf{k}\mu}$, $\bar{\alpha}_{\mathbf{k}\mu}$ and satisfying the commutation relation

$$\bar{\beta}_{\mathbf{k}\mu}\beta_{\mathbf{k}'\nu}-\beta_{\mathbf{k}'\nu}\bar{\beta}_{\mathbf{k}\mu} = -g_{\mu\nu}\,k_0\,\delta(k_1-k_1')\,\delta(k_2-k_2')\,\delta(k_3-k_3'). \tag{3·10}$$

Comparing (3·7) with (3·4) and (3·9) with (3·6) one finds, for $k_0 > 0$,

$$\left.\begin{aligned}\sqrt{2}\,\alpha_{\mathbf{k}\mu} &= \xi_{\mathbf{k}\mu} - \xi^*_{-\mathbf{k}\mu}, & \sqrt{2}\,\bar{\alpha}_{\mathbf{k}\mu} &= -\xi_{-\mathbf{k}\mu} + \xi^*_{\mathbf{k}\mu},\\ \sqrt{2}\,\beta_{\mathbf{k}\mu} &= -\xi_{-\mathbf{k}\mu} - \xi^*_{\mathbf{k}\mu}, & \sqrt{2}\,\bar{\beta}_{\mathbf{k}\mu} &= \xi_{\mathbf{k}\mu} + \xi^*_{-\mathbf{k}\mu}.\end{aligned}\right\} \tag{3·11}$$

It should be remembered here that the bar denotes the conjugate complex of a dynamical variable and the star denotes the adjoint of an operator, which are two different things in the present theory. From equations (3·11), one sees that $\bar{\alpha}_{\mathbf{k}\mu} = \alpha^*_{\mathbf{k}\mu}$ and $\bar{\beta}_{\mathbf{k}\mu} = -\beta^*_{\mathbf{k}\mu}$.

The wave function in the new theory is a function of the $\xi_{\mathbf{k}\mu}$ and $\xi_{-\mathbf{k}\mu}$ (keeping $k_0 > 0$), and in the old theory it is a function of the $\alpha_{\mathbf{k}\mu}$ and of certain variables of the **B** field, which may conveniently be taken to be the $\beta_{\mathbf{k}\mu}$. To set up the transformation connecting the new theory with the old, one must obtain the transformation function connecting the $\xi_{\mathbf{k}\mu}$, $\xi_{-\mathbf{k}\mu}$ representation with the $\alpha_{\mathbf{k}\mu}$, $\beta_{\mathbf{k}\mu}$ representation. For this purpose it is sufficient to deal with only one value for the vector **k** and one direction μ, since different values for these variables correspond to different degrees of freedom, for which the transformations are independent. It is convenient now to put $\xi_{-\mathbf{k}\mu} = \zeta_{\mathbf{k}\mu}$ and to drop the suffixes **k** and μ, so that equations (3·11) become

$$\left.\begin{aligned}\sqrt{2}\,\alpha &= \xi - \zeta^*, & \sqrt{2}\,\bar{\alpha} &= -\zeta + \xi^*,\\ \sqrt{2}\,\beta &= -\zeta - \xi^*, & \sqrt{2}\,\bar{\beta} &= \xi + \zeta^*.\end{aligned}\right\} \tag{3·12}$$

The value of **k** concerned may be regarded as one of a discrete set (by enclosing the electromagnetic field in a finite box, or by some other way), and the commutation relations (3·3), (3·8), (3·10) then become

$$\xi^*\xi - \xi\xi^* = c, \quad \zeta^*\zeta - \zeta\zeta^* = -c, \tag{3·13}$$

$$\bar{\alpha}\alpha - \alpha\bar{\alpha} = c, \quad \bar{\beta}\beta - \beta\bar{\beta} = c, \tag{3·14}$$

where, if we take the case of $\mu = 1$, 2 or 3, c is some positive number.

The ξ, ζ representation, or any representation of the type of the ξ representation introduced in Appendix II, is of a more general kind than the usual representations of quantum mechanics, since the variables ξ, ζ occurring in the wave functions do not correspond to self-adjoint operators. In dealing with such a representation one must therefore refer back to first principles. The ξ, ζ representation is defined so as to make the representative of ξ or ζ times any $|\rangle$ equal the representative of that $|\rangle$ multiplied by the ξ or ζ variable, respectively, thus

$$\langle \xi' \zeta' | \xi | \rangle = \xi' \langle \xi' \zeta' | \rangle, \quad \langle \xi' \zeta' | \zeta | \rangle = \zeta' \langle \xi' \zeta' | \rangle. \tag{3·15}$$

From the commutation relations (3·13) one can see that when ξ^* or ζ^* is multiplied into any $|\rangle$, its representative $\langle\xi'\zeta'|\rangle$ gets operated on by $c\,\partial/\partial\xi'$, $-c\,\partial/\partial\zeta'$, respectively, thus

$$\langle\xi'\zeta'|\xi^*|\rangle = c\frac{\partial}{\partial\xi'}\langle\xi'\zeta'|\rangle, \quad \langle\xi'\zeta'|\zeta^*|\rangle = -c\frac{\partial}{\partial\zeta'}\langle\xi'\zeta'|\rangle. \tag{3·16}$$

The conjugate complex of a representative $\langle\xi'\zeta'|\rangle$ is a function of the conjugate complex variables, $\xi^{*\prime}$, $\zeta^{*\prime}$ say, corresponding to the adjoint operators ξ^*, ζ^*, and should thus be written $\langle|\xi^{*\prime}\zeta^{*\prime}\rangle$. The conjugate complexes of equations (3·15) are therefore

$$\langle|\xi^*|\xi^{*\prime}\zeta^{*\prime}\rangle = \xi^{*\prime}\langle|\xi^{*\prime}\zeta^{*\prime}\rangle, \quad \langle|\zeta^*|\xi^{*\prime}\zeta^{*\prime}\rangle = \zeta^{*\prime}\langle|\xi^{*\prime}\zeta^{*\prime}\rangle, \tag{3·17}$$

and those of equations (3·16) are

$$\langle|\xi|\xi^{*\prime}\zeta^{*\prime}\rangle = c\frac{\partial}{\partial\xi^{*\prime}}\langle|\xi^{*\prime}\zeta^{*\prime}\rangle, \quad \langle|\zeta|\xi^{*\prime}\zeta^{*\prime}\rangle = -c\frac{\partial}{\partial\zeta^{*\prime}}\langle|\xi^{*\prime}\zeta^{*\prime}\rangle. \tag{3·18}$$

Let us now determine the form of the matrix $\langle\xi'\zeta'|\xi^{*\prime\prime}\zeta^{*\prime\prime}\rangle$. From (3·15)

$$\langle\xi'\zeta'|\xi|\xi^{*\prime\prime}\zeta^{*\prime\prime}\rangle = \xi'\langle\xi'\zeta'|\xi^{*\prime\prime}\zeta^{*\prime\prime}\rangle, \quad \langle\xi'\zeta'|\zeta|\xi^{*\prime\prime}\zeta^{*\prime\prime}\rangle = \zeta'\langle\xi'\zeta'|\xi^{*\prime\prime}\zeta^{*\prime\prime}\rangle,$$

and from (3·18)

$$\langle\xi'\zeta'|\xi|\xi^{*\prime\prime}\zeta^{*\prime\prime}\rangle = c\frac{\partial}{\partial\xi^{*\prime\prime}}\langle\xi'\zeta'|\xi^{*\prime\prime}\zeta^{*\prime\prime}\rangle,$$

$$\langle\xi'\zeta'|\zeta|\xi^{*\prime\prime}\zeta^{*\prime\prime}\rangle = -c\frac{\partial}{\partial\zeta^{*\prime\prime}}\langle\xi'\zeta'|\xi^{*\prime\prime}\zeta^{*\prime\prime}\rangle.$$

Hence

$$\left(c\frac{\partial}{\partial\xi^{*\prime\prime}} - \xi'\right)\langle\xi'\zeta'|\xi^{*\prime\prime}\zeta^{*\prime\prime}\rangle = 0, \qquad \left(c\frac{\partial}{\partial\zeta^{*\prime\prime}} + \zeta'\right)\langle\xi'\zeta'|\xi^{*\prime\prime}\zeta^{*\prime\prime}\rangle = 0.$$

Similarly, working from (3·16) and (3·17), one finds

$$\left(c\frac{\partial}{\partial\xi'} - \xi^{*\prime\prime}\right)\langle\xi'\zeta'|\xi^{*\prime\prime}\zeta^{*\prime\prime}\rangle = 0, \qquad \left(c\frac{\partial}{\partial\zeta'} + \zeta^{*\prime\prime}\right)\langle\xi'\zeta'|\xi^{*\prime\prime}\zeta^{*\prime\prime}\rangle = 0.$$

Hence
$$\langle\xi'\zeta'|\xi^{*\prime\prime}\zeta^{*\prime\prime}\rangle = \exp\left[(\xi'\xi^{*\prime\prime} - \zeta'\zeta^{*\prime\prime})/c\right], \tag{3·19}$$

apart from a constant factor, which may be taken to be unity without essential loss of generality.

The law of multiplication of representatives must now be chosen so as to make the $\langle\xi'\zeta'|\xi^{*\prime\prime}\zeta^{*\prime\prime}\rangle$ given by (3·19) play the part of the unit matrix.

It is easily seen that this condition is fulfilled if one takes the product of $\langle a|\xi^{*\prime}\zeta^{*\prime}\rangle$ and $\langle \xi'\zeta'|b\rangle$ to be

$$\langle a|b\rangle = S_{\xi'}S_{\zeta'}\langle a|\xi^{*\prime}\zeta^{*\prime}\rangle\langle\xi'\zeta'|b\rangle, \tag{3·20}$$

where the symbols $S_{\xi'}$, $S_{\zeta'}$ denote certain kinds of summation, which are defined as follows.

If $f(\xi^*)$ and $g(\xi)$ are two power series in ξ^* and ξ respectively, so that

$$f(\xi^*) = \Sigma_{n=0}^{\infty} f_n \xi^{*n}, \quad g(\xi) = \Sigma_{n=0}^{\infty} g_n \xi^n,$$

then

$$S_\xi f(\xi^*)\,g(\xi) = \Sigma_{n=0}^{\infty} c^n n!\, f_n g_n. \tag{3·21}$$

The definition of S_ζ is similar, with $-c$ substituted for c. These definitions lead immediately to

$$S_{\xi''}S_{\zeta''}\langle\xi'\zeta'|\xi^{*\prime\prime}\zeta^{*\prime\prime}\rangle\langle\xi''\zeta''|b\rangle = \langle\xi'\zeta'|b\rangle,$$
$$S_{\xi'}S_{\zeta'}\langle a|\xi^{*\prime}\zeta^{*\prime}\rangle\langle\xi'\zeta'|\xi^{*\prime\prime}\zeta^{*\prime\prime}\rangle = \langle a|\xi^{*\prime\prime}\zeta^{*\prime\prime}\rangle,$$

which are the required conditions for $\langle\xi'\zeta'|\xi^{*\prime\prime}\zeta^{*\prime\prime}\rangle$ to play the part of the unit matrix. (3·21) is equivalent to the following explicit definition for S_ξ,

$$S_\xi f(\xi^*)\,g(\xi) = (\pi c)^{-1}\iint_{-\infty}^{\infty} f(\xi^*)\,g(\xi)\,e^{-\xi^*\xi/c}\,d\xi_R\,d\xi_I, \tag{3·22}$$

where ξ_R and ξ_I denote the real and pure imaginary parts of the complex variable ξ. The corresponding definition for S_ζ is

$$S_\zeta f(\zeta^*)\,g(\zeta) = (\pi c)^{-1}\iint_{-\infty}^{\infty} f(\zeta^*)\,g(-\zeta)\,e^{-\zeta^*\zeta/c}\,d\zeta_R\,d\zeta_I. \tag{3·23}$$

The multiplication rule given by (3·20), (3·21) is connected with the physical interpretation of the wave functions $\langle\xi'\zeta'|\rangle$, namely that the probability of there having been m photons emitted into the momentum and energy state **k** (corresponding to the variable ξ) and n photons absorbed from this state is

$$c^m m!\,(-c)^n n!\,|\text{coefficient of } \xi'^m\zeta'^n \text{ in } \langle\xi'\zeta'|\rangle|^2. \tag{3·24}$$

this being the quantity whose sum over all m and n must be put equal to unity to normalize the wave function. The form (3·24) is a generalization of (2·4), arising on account of the c and $-c$ on the right-hand sides of the commutation relations (3·13). It gives a negative probability for an odd number of photons having been absorbed, in agreement with what is stated in the text.

Let us now determine the transformation function $\langle\alpha'\beta'|\xi^{*\prime}\zeta^{*\prime}\rangle$. From the first of equations (3·12)

$$\sqrt{2}\langle\alpha'\beta'|\alpha|\xi^{*\prime}\zeta^{*\prime}\rangle = \langle\alpha'\beta'|\xi-\zeta^{*}|\xi^{*\prime}\zeta^{*\prime}\rangle,$$

which gives, according to (3·18) and (3·17)

$$\sqrt{2}\,\alpha'\langle\alpha'\beta'|\xi^{*\prime}\zeta^{*\prime}\rangle = \left(c\frac{\partial}{\partial\xi^{*\prime}}-\zeta^{*\prime}\right)\langle\alpha'\beta'|\xi^{*\prime}\zeta^{*\prime}\rangle.$$

Similarly, the other three equations (3·12) give

$$\sqrt{2}\,c\,\frac{\partial}{\partial\alpha'}\langle\alpha'\beta'|\xi^{*\prime}\zeta^{*\prime}\rangle = \left(c\frac{\partial}{\partial\zeta^{*\prime}}+\xi^{*\prime}\right)\langle\alpha'\beta'|\xi^{*\prime}\zeta^{*\prime}\rangle,$$

$$\sqrt{2}\,\beta'\langle\alpha'\beta'|\xi^{*\prime}\zeta^{*\prime}\rangle = \left(c\frac{\partial}{\partial\zeta^{*\prime}}-\xi^{*\prime}\right)\langle\alpha'\beta'|\xi^{*\prime}\zeta^{*\prime}\rangle,$$

$$\sqrt{2}\,c\,\frac{\partial}{\partial\beta'}\langle\alpha'\beta'|\xi^{*\prime}\zeta^{*\prime}\rangle = \left(c\frac{\partial}{\partial\xi^{*\prime}}+\zeta^{*\prime}\right)\langle\alpha'\beta'|\xi^{*\prime}\zeta^{*\prime}\rangle.$$

The solution of these four equations is

$$\langle\alpha'\beta'|\xi^{*\prime}\zeta^{*\prime}\rangle = \sqrt{2}\exp\left[(\alpha'\beta'+\sqrt{2}\,\alpha'\xi^{*\prime}+\sqrt{2}\,\beta'\zeta^{*\prime}+\xi^{*\prime}\zeta^{*\prime})/c\right], \quad (3·25)$$

in which the arbitrary multiplying factor is chosen for convenience to be $\sqrt{2}$. A different choice would only mean a different weight factor in the α, β representation.

It remains to determine the law of multiplication for representatives in the α, β representation, which can best be done by first determining the unit matrix $\langle\alpha'\beta'|\alpha^{*\prime\prime}\beta^{*\prime\prime}\rangle$. We have

$$\begin{aligned}\langle\alpha'\beta'|\alpha^{*\prime\prime}\beta^{*\prime\prime}\rangle &= S_{\xi'}S_{\zeta'}\langle\alpha'\beta'|\xi^{*\prime}\zeta^{*\prime}\rangle\langle\xi'\zeta'|\alpha^{*\prime\prime}\beta^{*\prime\prime}\rangle\\ &= 2S_{\xi'}S_{\zeta'}\exp\left[(\alpha'\beta'+\sqrt{2}\,\alpha'\xi^{*\prime}+\sqrt{2}\,\beta'\zeta^{*\prime}+\xi^{*\prime}\zeta^{*\prime})/c\right]\\ &\qquad\times\exp\left[(\alpha^{*\prime\prime}\beta^{*\prime\prime}+\sqrt{2}\,\alpha^{*\prime\prime}\xi'+\sqrt{2}\,\beta^{*\prime\prime}\zeta'+\xi'\zeta')/c\right].\end{aligned}$$

Performing the operation $S_{\xi'}$ according to (3·22), one gets

$$\begin{aligned}&\langle\alpha'\beta'|\alpha^{*\prime\prime}\beta^{*\prime\prime}\rangle\\ &\quad= 2S_{\zeta'}\exp\left[(\alpha'\beta'+\sqrt{2}\,\beta'\zeta^{*\prime}+\sqrt{2}\,\beta^{*\prime\prime}\zeta'+\alpha^{*\prime\prime}\beta^{*\prime\prime})/c\right]\\ &\qquad\times(\pi c)^{-1}\iint\exp\{[(\sqrt{2}\alpha'+\zeta^{*\prime})\xi^{*\prime}+(\sqrt{2}\,\alpha^{*\prime\prime}+\zeta')\xi'-\xi^{*\prime}\xi']/c\}\,d\xi'_R\,d\xi'_I\\ &\quad= 2S_{\zeta'}\exp\left[(\alpha'\beta'+\sqrt{2}\,\beta'\zeta^{*\prime}+\sqrt{2}\,\beta^{*\prime\prime}\zeta'+\alpha^{*\prime\prime}\beta^{*\prime\prime})/c\right]\\ &\qquad\times\exp\left[(\sqrt{2}\,\alpha'+\zeta^{*\prime})(\sqrt{2}\,\alpha^{*\prime\prime}+\zeta')/c\right].\end{aligned}$$

Again, performing the operation $S_{\zeta'}$ according to (3·23), one gets

$$\begin{aligned}
&\langle\alpha'\,\beta'|\alpha^{*''}\,\beta^{*''}\rangle \\
&\quad = 2\exp\left[(\alpha'\,\beta'+\alpha^{*''}\,\beta^{*''}+2\alpha'\,\alpha^{*''})/c\right] \\
&\qquad \times(\pi c)^{-1}\iint\exp\{[\sqrt{2}\,(\beta'+\alpha^{*''})\,\zeta^{*'}-\sqrt{2}\,(\beta^{*''}+\alpha')\,\zeta'-2\zeta^{*'}\,\zeta']/c\}\,d\zeta'_R\,d\zeta'_I \\
&\quad = \exp\left[(\alpha'\,\beta'+\alpha^{*''}\,\beta^{*''}+2\alpha'\,\alpha^{*''})/c\right]\exp\left[-(\beta'+\alpha^{*''})\,(\beta^{*''}+\alpha')/c\right] \\
&\quad = \exp\left[(\alpha'\,\alpha^{*''}-\beta'\,\beta^{*''})/c\right]. \qquad (3{\cdot}26)
\end{aligned}$$

This result is similar to (3·19). It shows that the product of a wave function $\langle\alpha'\,\beta'|b\rangle$ and a conjugate complex wave function $\langle a|\alpha^{*'}\,\beta^{*'}\rangle$ is, corresponding to (3·20),

$$\langle a|b\rangle = S_{\alpha'}\,S_{\beta'}\,\langle a|\alpha^{*'}\,\beta^{*'}\rangle\,\langle\alpha'\,\beta'|b\rangle, \qquad (3{\cdot}27)$$

where S_α is defined like S_ξ in (3·21) and S_β is defined like S_ζ.

The multiplication rule (3·27) leads to the physical interpretation of the wave functions $\langle\alpha'\,\beta'|\rangle$. It gives, for the probability of there being m photons of the **A** field and n photons of the **B** field in the momentum and energy state **k**, the result

$$c^m\,m!\,(-c)^n\,n!\,|\text{coefficient of } \alpha'^m\,\beta'^n \text{ in } \langle\alpha'\,\beta'|\rangle|^2, \qquad (3{\cdot}28)$$

analogous to (3·24). There is now complete mathematical symmetry between the pair of variables α, β and the pair ξ, ζ, though there remains the difference of physical interpretation, that α and β refer to the existence of photons in two fields and ξ and ζ refer to the emission and absorption of photons.

Let us apply the transformation to the initial state of the field, for which, of course, no emissions or absorptions have taken place, so that the ξ', ζ' wave function is

$$\langle\xi'\,\zeta'|\text{init}\rangle = 1. \qquad (3{\cdot}29)$$

Then $\langle\alpha'\,\beta'|\text{init}\rangle = S_{\xi'}\,S_{\zeta'}\,\langle\alpha'\,\beta'|\xi^{*'}\,\zeta^{*'}\rangle\,\langle\xi'\,\zeta'|\text{init}\rangle$

$$= \sqrt{2}\,S_{\xi'}\,S_{\zeta'}\exp\left[(\alpha'\,\beta'+\sqrt{2}\,\alpha'\,\xi^{*'}+\sqrt{2}\,\beta'\,\zeta^{*'}+\xi^{*'}\,\zeta^{*'})/c\right]$$

from (3·25). The operators $S_{\xi'}, S_{\zeta'}$ now reduce to putting $\xi^{*'} = 0$, $\zeta^{*'} = 0$ and hence

$$\langle\alpha'\,\beta'|\text{init}\rangle = \sqrt{2}\,e^{\alpha'\beta'/c}. \qquad (3{\cdot}30)$$

Applying the interpretation (3·28), one finds, firstly, that the probability of there being m photons of the **A** field and n of the **B** field in the state **k** is zero unless $m = n$. This is what one would expect from the assumption that the **B** field equals the **A** field initially, an assumption that was made in the text. Secondly, one finds that the probability of there being n photons of the **A** or the **B** field in the state **k** is $P_n = (-)^n 2$, a result which is used in the text.

References

Dirac, P. A. M. 1938 *Proc. Roy. Soc.* A, **167**, 148–169.
Dirac, P. A. M. 1939*a* *Proc. Camb. Phil. Soc.* **35**, 416–418.
Dirac, P. A. M. 1939*b* *Ann. Inst. Poincaré.* This paper has not yet appeared in England.
Fock, V. 1934 *Phys. Z. Sowjet*, **6**, 425–469.
Fokker, A. D. 1929 *Z. Phys.* **58**, 386–393.
Gordon, W. 1926 *Z. Phys.* **40**, 117–133.
Heisenberg, W. & Pauli, W. 1929*a* *Z. Phys.* **56**, 1–61.
Heisenberg, W. & Pauli, W. 1929*b* *Z. Phys.* **59**, 168–190.
Jordan, P. & Pauli, W. 1928 *Z. Phys.* **47**, 151–173.
Klein, O. 1927 *Z. Phys.* **41**, 407–442.
Pauli, W. & Weisskopf, V. 1934 *Helv. Phys. Acta*, **7**, 709–731.
Weiss, P. 1936 *Proc. Roy. Soc.* A, **156**, 192–220.

Molecular structure and rubber-like elasticity

I. The crystal structures of β gutta-percha, rubber and polychloroprene

By C. W. Bunn, B.A., B.Sc.

The Research Department of Imperial Chemical Industries Ltd., Northwich

(*Communicated by Sir William Bragg, F.R.S.—Received* 13 *March* 1941)

The crystal structures of β gutta-percha, rubber and polychloroprene have been determined by interpretation of X-ray diffraction photographs.

β Gutta-percha ($-CH_2-C(CH_3)=CH-CH_2-)_n$ is orthorhombic, with axial lengths $a_0 = 7{\cdot}78$ A, $b_0 = 11{\cdot}78$ A, $c_0 = 4{\cdot}72$ A. Four long-chain molecules pass through this cell parallel to the c axis. The space group is $P\,2_12_12_1$, and the co-ordinates of the five carbon atoms of one structural unit are:

(CH_2)	(C)	(CH)	(CH_2)	(CH_3)
0·926	0·000	0·000	0·074	0·970
0·110	0·146	0·074	0·110	0·277
0·676	0·960	0·177	0·462	0·980

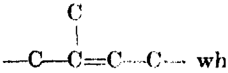
The molecules are asymmetric; all the molecules passing through any one crystal are identical—either left-handed or right-handed, not mixed. The carbon chain is a non-planar zigzag; each chain unit $C-C=C-C$ is planar, and has the *trans* configuration, but the connecting links (CH_2-CH_2) lie in a different plane; the plane $-C-C-C=$ makes an angle of 115° with plane $-C-C=C-$.

α Gutta-percha molecules probably also have the *trans* double-bond configuration, and differ from β molecules only in the positions of the $CH_2—CH_2$ bonds; the chain form is thus different from that of β molecules.

Crystalline rubber (also $(—CH_2—C(CH_3)=CH—CH_2—)_n$) is monoclinic, and has $a_0 = 12·46$ A, $b_0 = 8·89$ A, $c_0 = 8·10$ A, $\beta = 92°$. Four long-chain molecules pass through this cell parallel to the *c* axis. The space group is $P\,2_1/a$, and the co-ordinates of the ten carbon atoms of one structural chain unit are

(CH_2)	0·753	(C)	0·854	(CH)	0·845	(CH_2)	0·745	(CH_3)	0·968
	0·899		0·865		0·905		0·959		0·876
	0·802		0·703		0·542		0·457		0·773
(CH_2)	0·744	(C)	0·644	(CH)	0·659	(CH_2)	0·757	(CH_3)	0·532
	0·834		0·874		0·905		0·834		0·828
	0·326		0·215		0·052		0·975		0·268

The molecules are asymmetric; two of the molecules passing through the unit cell are left-handed and two right-handed. The chain carbon atoms of any one molecule form a non-planar zigzag, in which the double-bond units C—C=C—C have the *cis* configuration. The two isoprene units



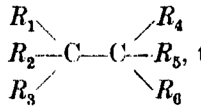
$—C—\overset{\overset{C}{|}}{C}=C—C—$ which make up the structural unit are not identical in configuration; the distortion appears to be due to intermolecular forces.

The structure of crystalline polychloroprene $(—CH_2—CCl=CH—CH_2—)_n$ is completely analogous to that of β gutta-percha. The cell is orthorhombic, with $a_0 = 8·84$ A, $b_0 = 10·24$ A, $c_0 = 4·79$ A. The co-ordinates of carbon and chlorine atoms are

(CH_2)	0·077	(C)	0·033	(CH)	0·008	(CH_2)	0·970	Cl	0·167
	0·054		0·115		0·034		0·099		0·243
	0·278		0·000		0·787		0·509		0·000

INTRODUCTION

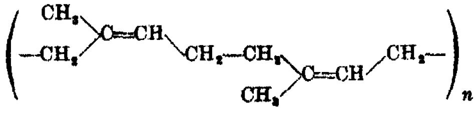
The extraordinary mechanical properties of rubber, which have led to its widespread use for a variety of special purposes, seemed until recently very puzzling. Rubber did not appear to fit into any of the simple categories recognized by physical chemistry in the last century; in the ordinary condition at room temperature it is not a crystalline solid, nor a liquid; it could perhaps be classed with the glasses as a non-crystalline solid, but its softness, flexibility and above all its enormous reversible extensibility, separate it widely from most other non-crystalline solids; its peculiar properties suggest a semi-liquid nature; or perhaps it might be described as a viscous liquid having almost the coherence of a solid. Its mechanical properties are more like those of animal tissues than those of most laboratory chemicals, and the possibility that this similarity may not be entirely superficial lends additional interest to the search for an explanation of these mechanical properties in terms of molecular physics.

Within the last few years the possibility of arriving at such an explanation seems to have been brought nearer by the development of two conceptions. The first is that rubber is a long-chain polymer. It was known that rubber is a polymer of isoprene $CH_2{=}C(CH_3){-}CH{=}CH_2$, and chemical evidence shows that the isoprene units are linked end to end (1 : 4) (Harries 1905, 1912). The viscosities of rubber solutions suggest strongly that the molecules are very long chains (Staudinger 1930). Chemically, raw rubber thus appears to be $(-CH_2-C(CH_3){=}CH-CH_2-CH_2-C(CH_3){=}CH-CH_2-)_n$. The second conception is that such long-chain molecules are likely to be very flexible. When two carbon atoms are joined together by a single bond:  $\overset{R_1}{\underset{R_3}{R_2-C}}-\overset{R_4}{\underset{R_6}{C-R_5}}$, the left-hand group can usually rotate with respect to the right-hand group, round the central connecting link as axis. Such rotation round each of the single bonds of a long-chain molecule would obviously give rise to great flexibility; the molecule could easily assume an enormous number of different geometrical forms. It is natural to suppose that the flexibility of rubber is in some way due to the flexibility of the molecules themselves. Thus Mack (1934) pictures the molecules in ordinary rubber as being closely coiled up, while Mark (1939), Kuhn (1936, 1939) and Guth (1937) suppose them to have a quite irregular configuration; the stretching of rubber is thought to be due to the straightening out of the molecules.

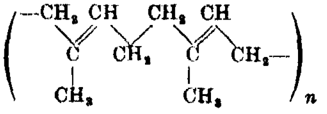
Not all long-chain polymers have rubber-like properties at room temperature, however. Consider, for instance, the properties of gutta-percha. This substance is, like rubber, a 1 : 4 chain-polymer of isoprene (Harries 1905, 1912), yet its properties at room temperature are quite different from those of rubber; it is a crystalline solid, tough and horn-like in character. Thus, there are two substances, both chain polymers of the same monomer, which have quite different mechanical properties at room temperature. It is true that the average length of gutta-percha molecules (about 3000 carbon atoms) is apparently only about one-third that of raw rubber molecules (about 10,000 carbon atoms); but the difference in their mechanical properties is not solely or even mainly due to the difference in the lengths of the molecules, since rubber degraded to about the same molecular length as gutta-percha is still rubber-like and non-crystalline. If rubber-like properties are due to molecular flexibility, it appears that gutta-percha molecules are not flexible at room temperature; the supposition being that there is some geometrical difference between rubber and gutta-percha molecules, such that the latter have not the flexibility of the former.

Rubber in the ordinary condition at room temperature is amorphous, but Katz (1925) discovered that it becomes crystalline on stretching to several times its original length. When a stretched specimen is released, the rapid contraction is accompanied by the disappearance of the crystalline structure and a return to the amorphous condition. The extraordinary mechanical properties of the substance are thus paralleled by equally extraordinary behaviour of the molecules, which actually arrange themselves in a precise pattern during stretching and disarrange themselves on release. It is natural to suspect a vital connexion between the two sets of phenomena, and to suppose that if it were known why the molecules were able to arrange and disarrange themselves in response to mere mechanical changes, it might then also be understood why the substance was elastic. The geometrical structure of the molecules themselves, and perhaps the nature of the arrangement of them in the crystal, should suggest some reason for such remarkable behaviour.

The X-ray diffraction photographs of crystalline rubber and gutta-percha show at once that the molecules are geometrically different. When rubber is stretched, it is natural to suppose that the long molecules become parallel to the direction of stretching, and the identity period of the crystals along the direction of stretching is therefore the identity period of the molecules themselves; this identity period is found to be 8·1 A (Mark & Susich 1928). Gutta-percha, being already crystalline at room temperature, cannot be stretched reversibly to this extent, but it can be drawn out or rolled out irreversibly* so that all the crystals become parallel to the direction of drawing or rolling. Assuming that the long molecules are parallel to the direction of drawing or rolling, the identity period of the crystals along this direction is the identity period of the molecules. There are two crystalline forms of gutta-percha, known as α and β, and their identity periods in the direction of drawing are 8·9 and 4·7 A respectively (Hauser & Susich 1931). It appears, therefore, that there are two different chain forms of gutta-percha, both different from that of rubber. An obvious possibility of isomerism in 1 : 4 polymers of isoprene arises from the presence of double bonds in the chain, and it has been suggested (Meyer & Mark 1928) that rubber is the *cis* form, with chain bonds on the same side of the double bond, while gutta-percha is the *trans* form, with chain bonds on opposite sides of the double bond (Fuller 1936):



cis



trans

* The process becomes reversible at a temperature just below the melting-point.

If this is true, the question arises, 'Why is the *cis* form of chain more flexible than the *trans* form?' To answer this question, it is necessary to know more about the precise geometry of the molecules.

The position is complicated by the existence of two forms of gutta-percha. These two forms, which can be easily converted into each other but not into rubber, are evidently stereoisomers having different chain forms; if both have the *trans* double-bond configuration, they must differ in single-bond configuration.

Another substance whose molecular geometry appears to be of crucial importance is polychloroprene, the synthetic rubber-like substance which is the basis of 'Neoprene' synthetic rubbers. Chemically, that is a 1 : 4 polymer, and differs from rubber and gutta-percha in having chlorine atoms in place of substituent methyl groups:

$$(\text{—CH}_2\text{—CCl=CH—CH}_2\text{—CH}_2\text{—CCl=CH—CH}_2\text{—})_n$$

(Carothers, Williams, Collins & Kirby 1931). The remarkable thing about this substance is that its identity period (4·79 A) is almost exactly the same as that of β gutta-percha (4·72 A), suggesting the same chain structure, yet its mechanical properties at room temperature are like those of rubber, not those of gutta-percha. If rubber-like properties are due to molecular flexibility, why does the substitution of chlorine atoms in place of methyl groups increase the flexibility of the molecules?

The answers to such questions are presumably to be found in the geometry of the molecules. The first part of the present work has therefore been the elucidation of the crystal structures of the three substances mentioned—rubber, gutta-percha and polychloroprene. The β form of gutta-percha was studied first, as it appeared to offer the greatest chances of success; its identity period (4·7 A) is shorter than that of rubber (8·1 A), indicating a simpler chain structure, and it gives more detailed X-ray photographs than polychloroprene.

(*a*) β GUTTA-PERCHA

The two crystalline forms of gutta-percha

The crystalline form which occurs in the tree is known as the α form. If heated above 65° C it becomes transparent, optically isotropic, and amorphous, giving an X-ray diffraction pattern of diffuse rings. If this amorphous material is cooled very slowly (0·5° C per hr.) the α form recrystallizes, but if cooling is rapid the β form crystallizes. The latter, on heating, becomes amorphous at 56° C. Thus the β form, since it has the lower crystallization point and is the first to crystallize on rapid cooling, is strictly speaking the less stable of the two at all temperatures; however, it appears to be permanently metastable at room temperature (Hauser & Susich 1931).

The X-ray diffraction photographs of the α and β forms are quite different, and the identity periods of drawn or rolled specimens of the two forms are different. Evidently this is a case of stereoisomerism, not one of the frequent examples of two different packings of the same molecules; the chain forms of the molecules are different. The two types of molecules are easily interconvertible, but neither of the forms can be converted into rubber.

Chain form of β molecules

The identity period of β gutta-percha (4·7 A) is so short that it is probable from the start that there is only one isoprene unit in this length. Of the *cis* and *trans* forms of molecule, only the latter is likely to have an identity period with one isoprene unit in it. A *trans* chain structure with all the carbon atoms in the same plane (figure 1 *c*, *d*) has been suggested for β gutta-percha (Fuller 1936). But accepting Robertson's interatomic distances (1938) (C—C 1·54 A, C=C 1·33 A) and the theoretical bond angles ($>$C 109½°, $\gg$C 125°), and assuming that the $\diagup^{CH_2}\diagdown_{CH_2}\diagup$ group forms a plane zigzag as in the long-chain normal paraffins (Müller 1928; Hengstenberg 1928; Bunn 1939), a planar *trans* chain (figure 1*c*, *d*) would have an identity period of 5·04 A, a figure considerably in excess of the actual period of β molecules (4·7 A). The only way of attaining the correct identity period without serious and improbable alterations of the interatomic distances and bond angles is to assume a non-planar chain structure. Now, in ethylene (Thompson 1939) and its chloro-substitution products (Brockway, Beach & Pauling 1935), all the atoms in the grouping ${}^{a}_{b}\!>\!C{=}C\!<^{c}_{d}$ are in the same plane. It is expected, therefore, that atoms 1, 2, 3 and 4 of the chain (figure 1*c*), as well as the methyl carbon 5, will all be in the same plane. Each isoprene unit is thus expected to be planar, as far as its carbon atoms are concerned. The only way of attaining the correct identity period of 4·7 A while keeping each isoprene unit planar is to move bond 4–1*a* out of the plane 12345 by rotation round bond 3–4; at the same time, in order

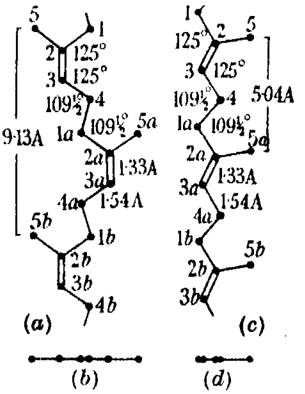


FIGURE 1. Planar poly-isoprene models. (*a*) *Cis* chain. (*b*) End view of (*a*). (*c*) *Trans* chain. (*d*) End view of (*c*).

to keep isoprene unit 1a2a3a4a5a strictly parallel to isoprene unit 12345 so that the two remain crystallographically equivalent, it is necessary to rotate bond 1a–4 round bond 1a–2a. Thus one isoprene unit has been moved towards the other along the chain axis while maintaining the correct distance between atoms 4 and 1a and maintaining the angles 341a and 41a2a at 109½°. There are two ways of doing this; bond 4–1a can be rotated either clockwise or anti-clockwise; two different types of asymmetric molecule, each the mirror image of the other, are obtained in this way; they are shown in figure 2. The end-views of such molecules show the four chain carbon atoms of each isoprene unit at the corners of a parallelogram. If each isoprene unit is planar, these are the only possible *trans* chain structures having a one-unit period of 4·7 A.

Crystal structure determination

X-ray diffraction photographs of a piece of cold-rolled sheet were taken on a cylindrical camera with Cu $K\alpha$ radiation. Rolling has the advantage over drawing that it produces a double orientation effect; not only does the *c* axis of each crystal become nearly parallel to the direction of rolling, but also the *a* axis tends to lie in the plane of the sheet; this lateral orientation is far less perfect than the longitudinal orientation, but is useful for the purpose of checking the indices of the diffraction spots. In addition to the usual 'fibre photographs' (with specimen vertical), another photograph was taken with the specimen horizontal and oscillating from a position normal to the X-ray beam to a position 45° to the beam; this photograph recorded the diffracted beams from the crystal plane perpendicular to the chain axis (the *c* axis). Similar oscillation photographs of rubber and polychloroprene specimens were subsequently taken during the work on those substances.

All the diffracted beams fit an orthorhombic cell containing four isoprene units. The photographs agree with those of Fuller (1936) and the electron diffraction photographs of Storks (1938), except in one particular; Fuller records some weak spots which appear to demand a larger cell; since these were not found even in dense photographs (and are absent from Storks's photographs), they must have been due to impurities in Fuller's specimens.

The *c* projection, giving an end-view of the molecules, was considered first. The problem was simplified by an apparent halving of the *b* axis from this viewpoint, a feature shown later to be an 'accidental' pseudo-symmetry, not a significant halving due to space group symmetry. It was found that no arrangement of planar molecules satisfies the (visually estimated) *hk*0 intensities; the only atomic positions which are satisfactory indicate a non-planar molecule whose end-view agrees approximately with that anticipated in figure 2.

The space group was considered next. The possibilities are limited by the following consideration. β Gutta-percha is made by cooling amorphous ('melted') material rapidly. 'Melted' gutta-percha is not a liquid but a somewhat rubber-like solid, hence the molecules do not move about relative to each other to any great extent. (If they did, the material would be fluid.) Neither can these enormously long molecules turn round to reverse their ends. Hence, on crystallization, the molecules settle down in an orderly manner while remaining more or less where they happen to be. There are evidently bundles of roughly parallel molecules in the amorphous material; all that happens during crystallization is that the molecules arrange themselves precisely without moving about bodily to any great extent. Since

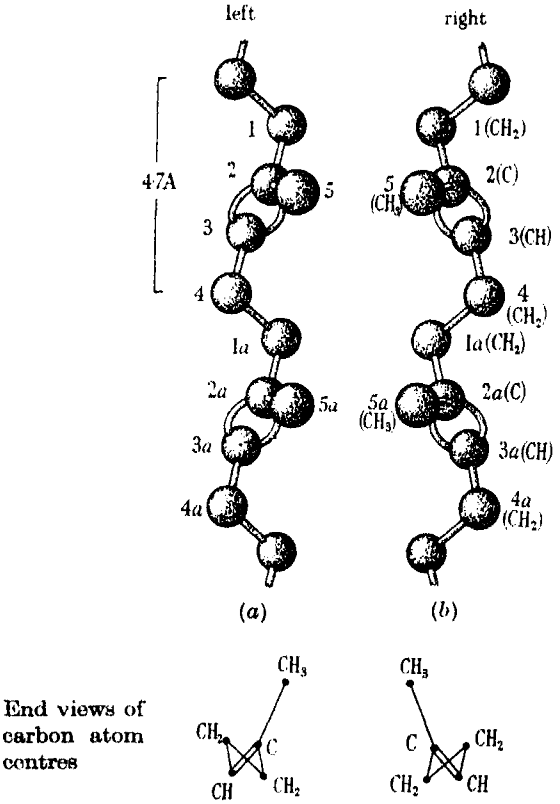


FIGURE 2. Molecular models assuming planar isoprene units (*trans*) and identity period of 4·7 A. Left- and right-handed molecules.

in any bundle of molecules there are likely to be approximately equal numbers of 'up' and 'down' molecules (an 'up' molecule being defined as one with the methyl group above the double bond, and a 'down' molecule the reverse), it is probable that each crystal contains equal numbers of 'up' and 'down' molecules; thus, of the four molecules passing through the unit cell, two are likely to be 'up' and two 'down'. Accepting this conclusion, and accepting also the approximate molecular positions found in the preliminary calculations of $hk0$ intensities, only three space groups are possible: $P\,2_12_12_1$, $P\,2_12_12$, and Pca (the last with axes interchanged $abc \rightarrow cab$). $P\,2_12_12$ is ruled out by the absence of odd orders of $00l$. The final choice of space group was made, and atomic positions (for carbon atoms) found, by consideration of the intensities of all the diffraction spots on the photographs. A fairly good agreement between observed and calculated intensities was obtained for the $P\,2_12_12_1$ arrangement.* The calculations were much shortened by using charts of the type devised by Bragg & Lipson (1936). The method of using these for hkl intensities is described in an appendix.

Description of crystal structure

The unit cell is orthorhombic, with $a_0 = 7{\cdot}78$ A, $b_0 = 11{\cdot}78$ A, $c_0 = 4{\cdot}72$ A. Four long-chain molecules pass through this cell parallel to the c axis. The space group is $P\,2_12_12_1$, and the co-ordinates xyz of the carbon atoms of one right-handed 'down' molecule ('right-handed' is defined in figure 2) are the following:

C_1	0·926*a*	C_2	0·000*a*	C_3	0·000*a*	C_4	0·074*a*	C_5	0·970*a*	x
(CH_2)	0·110*b*	(C)	0·146*b*	(CH)	0·074*b*	(CH_2)	0·110*b*	(CH_3)	0·277*b*	y
	0·676*c*		0·960*c*		0·177*c*		0·462*c*		0·980*c*	z

The atoms of the other three molecules in the unit cell are at the other equivalent positions of space group $P\,2_12_12_1$, i.e.

$$\tfrac{1}{2}-x,\ -y,\ \tfrac{1}{2}+z; \quad \tfrac{1}{2}+x,\ \tfrac{1}{2}-y,\ -z; \quad -x,\ \tfrac{1}{2}+y,\ \tfrac{1}{2}-z.$$

All four molecules in the unit cell are identical—all right-handed in the example given here and illustrated by figure 3*a*, *b*; the crystal may be referred to as a right-handed crystal. Presumably, in any specimen, there are equal proportions of left- and right-handed crystals. The fact that in any one crystal all the molecules are either left-handed or right-handed, not mixed, attains some importance in the discussion of the origin of elasticity (Part III of this work).

---

* The best agreement between calculated and observed intensities was obtained by assuming that the diffracting powers of C, CH, CH_2 and CH_3 are in the ratios 6:7:8:9 (the ratios of the number of electrons in the groups).

The structure of individual molecules differs very little from the structure predicted in figure 2. The chain-carbon atoms of each isoprene unit lie in a plane; atoms 1 and 4 are on opposite sides of the double bond, the chain thus being the *trans* form. All interatomic distances and bond angles in the chain are normal*: C—C 1·54 A, C=C 1·33 A, >C 109½°, >C 125°.

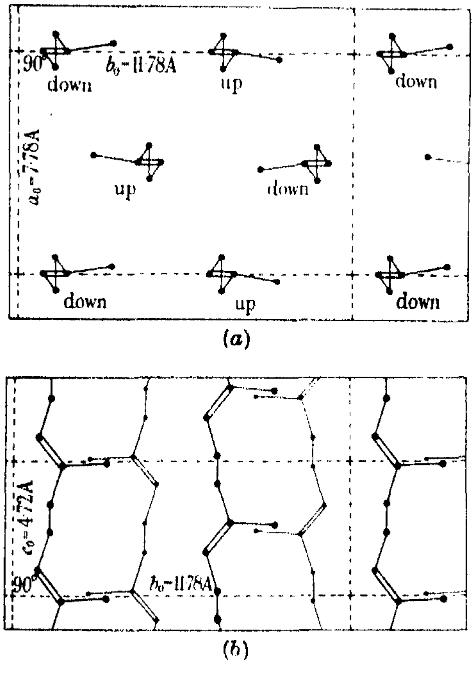


FIGURE 3. Structure of right-handed β gutta-percha crystal, seen (*a*) along *c* axis, (*b*) along *a* axis.

The more noteworthy features of the molecular structure are illustrated more clearly in figure 4. The most important feature is the non-planar character of the chain; bond 1–4*b*, if it were in plane 1234, would lie along the broken line in the diagram; actually it makes an angle of 63° with this line; bond 4–1*a* is at the same angle on the opposite side of plane 1234. (Expressed otherwise, plane 4*b*–1–2 makes an angle of 115° with plane 1234.) The reason for this non-planar character is discussed, in connexion with the stereochemistry of other aliphatic molecules, in Part II of this work.

---

* The procedure adopted in the calculations was, first of all to assume normal interatomic distances and bond angles, and to make alterations only if demanded by the intensities. For the chain carbon atoms no alterations seemed necessary.

The position of the methyl group needs comment. The carbon atom of this group (5) is not in plane 1234; it is displaced 0·65 A away from its 'ideal' position in plane 1234. The main cause of this distortion is evidently the repulsion of the chain CH_2 group 4*b* (figure 4). The effective radius of a CH_2 or CH_3 group is 1·80–2·05 A (depending on the direction—these groups are not spherical), and in all structures so far determined, the distance between such groups in neighbouring molecules is 3·6–4·1 A. The distance between 4*b* (a CH_2 group) and 5 (the CH_3 group) even in their present positions is only 3·2 A; if the methyl group were in its 'ideal' position in plane 1234, the distance would be still smaller. The reason for the distortion is thus obvious, but it remains surprising that it is so large. The determined position, however, must be accepted; although no great accuracy can be claimed, yet it seems quite impossible to account for the X-ray intensities except by assuming that the CH_3 group is in the position already given (within ± 0·1 A along *a*). Distortions of similar magnitude are found in the structure of rubber; and here it appears that intermolecular forces may also cause distortions (see below).

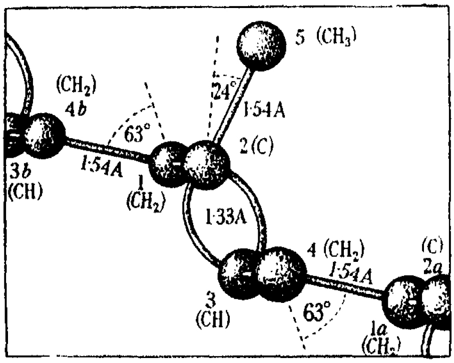
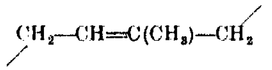


FIGURE 4. Details of structure of left-handed β gutta-percha molecule. Distances 1–2 and 3–4 = 1·54 A. Angles 4*b*–1–2 and 3–4–1*a* = 109½°. Broken straight lines are in the plane 1234.

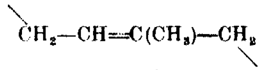
Probable structure of α gutta-percha molecules

Knowledge of the structure of β gutta-percha appears to provide the key to the molecular structures of α gutta-percha and of rubber, for, assuming bond orientations as in β gutta-percha, it is possible to construct chains having the correct identity periods for the other two substances.

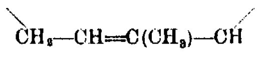
The α form of gutta-percha presumably has a *trans* double-bond configuration like the β form but differently oriented connecting links between the *trans* isoprene units. The reason for the non-planar character of the β chain will be discussed later, in Part II; here it will merely be assumed that the CH_2—CH_2 chain bonds in other molecules make the same angle with the plane of the isoprene unit as in β gutta-percha molecules. Four different orientations of the two connecting links of any one isoprene unit are possible; they can be represented (as well as is possible on a flat surface) thus:



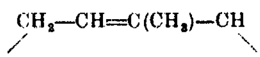
1



2



3



4

The plane of the isoprene unit is supposed to be perpendicular to the plane of the paper, and the connecting links nearly in the plane of the paper. The methyl groups are supposed to be above the plane of the paper.

If a large number of units of type 1 are joined together, a left-handed β gutta-percha molecule (figure 2*a*) is obtained; if units of type 2 are used, a right-handed β molecule (figure 2*b*) is the result. Units of type 3 alone, or 4 alone, give complex spiral molecules with very long periods, but if 3 and 4 are joined together alternately, a molecule with a two-unit period of 8·9 A is obtained (figure 5*a*). 8·9 A is exactly the known identity period of α gutta-percha, hence figure 5*a* probably represents the structure of α molecules.

There is one other type of molecule with a two-unit period, obtained by joining types 1 and 2 together alternately (figure 5*b*). The identity period of this molecule, which is thus a third possible form (γ) of gutta-percha, is 9·4 A, but so far such a form has not been encountered. Other combinations of the above four types of units give more complex molecules with very long periods, but these also have not been encountered, the only known forms of gutta-percha being the two already considered. (The tendency in nature is, of course, towards simplicity wherever possible.)

(*b*) RUBBER

Possible cis *chain forms*

Since gutta-percha has the *trans* double-bond configuration, rubber is presumably the *cis* form. The identity period of rubber molecules in a crystal is 8·1 A. A *cis* chain with all the carbon atoms in the same plane (figure 1 *a*, *b*) would have an identity period of 9·13 A, assuming the normal interatomic distances and bond angles. Therefore it is to be expected that the carbon chain in rubber molecules is non-planar. Each isoprene chain unit is expected to be substantially planar, hence likely non-planar chain forms can only be constructed by placing the CH_2—CH_2 connecting links out of the planes of the isoprene units. It is reasonable to suppose that the angles between the connecting links and the planes of the isoprene units would be similar to those in gutta-percha molecules. It has been seen that

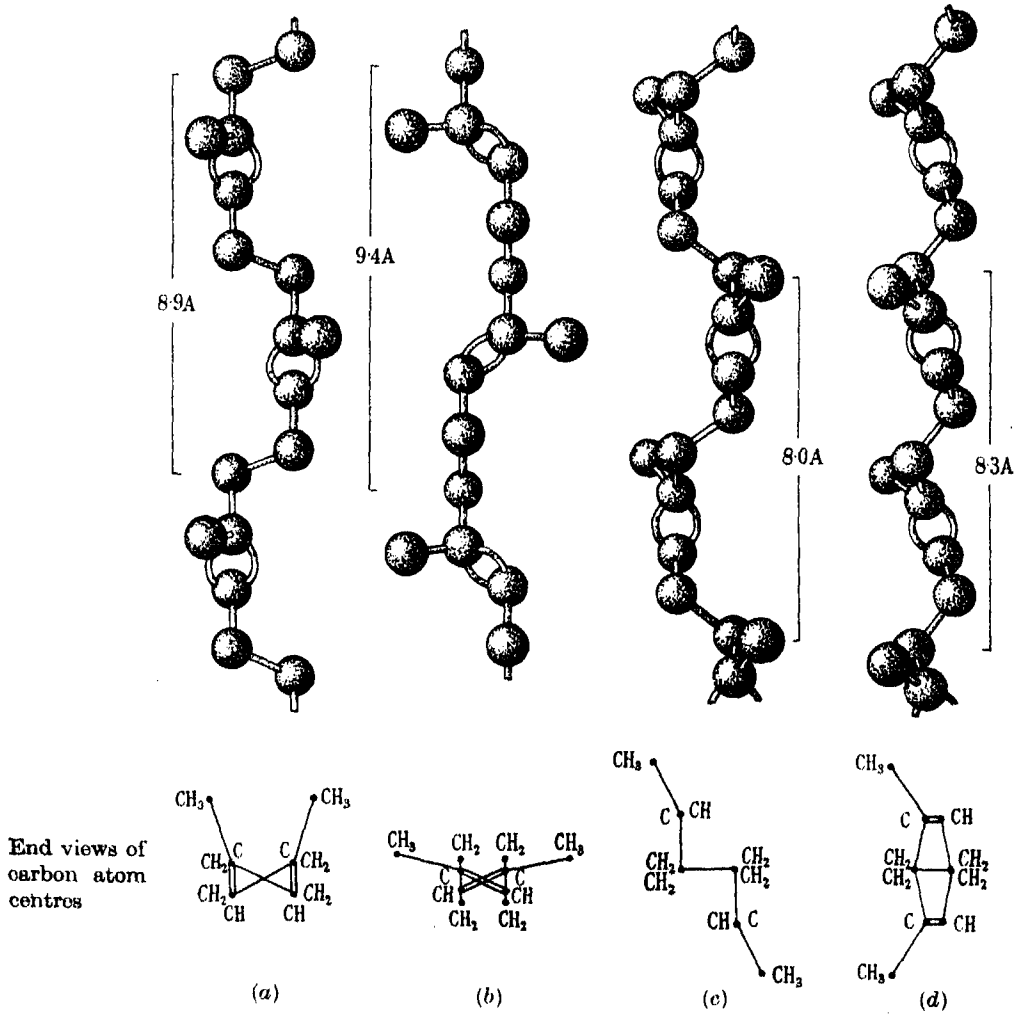


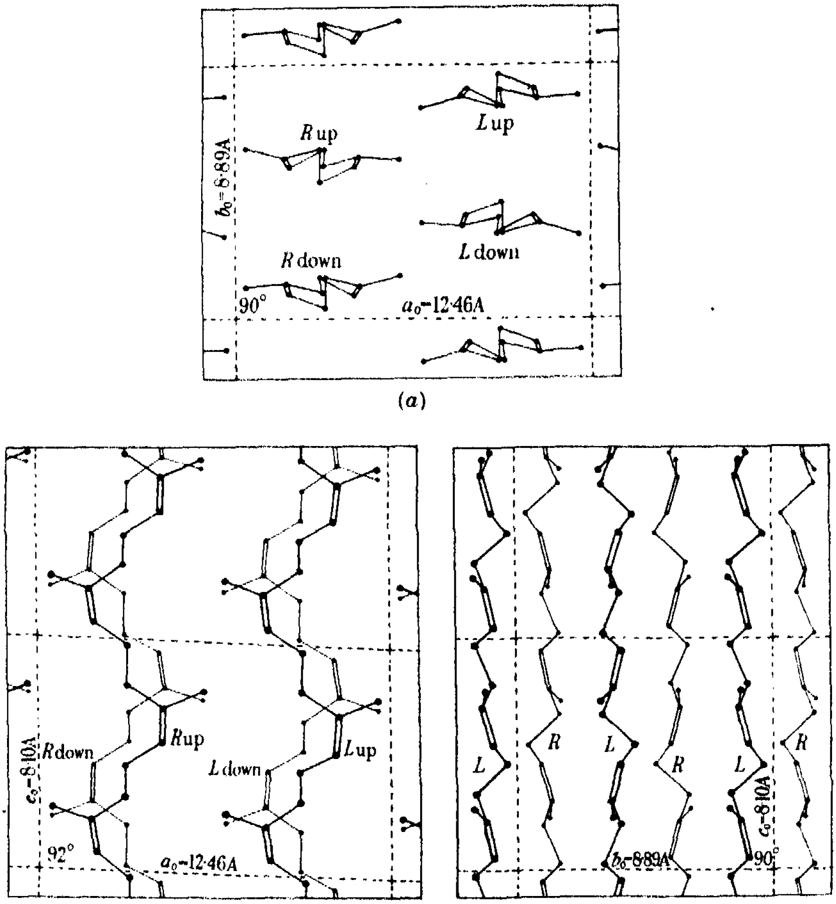
FIGURE 5. (*a*) Probable structure of gutta-percha. (*b*) Third possible form of gutta-percha (unknown), (*c*) and (*d*) *Cis* chains constructed according to the same principles of single-bond orientation.

there are four different arrangements of the pair of connecting links of each isoprene unit. Many different chain forms could be constructed by using these arrangements in different combinations, but it is found that there is no *cis* model with a one-unit period, and only two possible types with two-unit periods. One of these (figure 5*c*), a molecule with a period of about 8·0 A, is constructed exclusively of units of type 3 (left-handed molecule), or exclusively of type 4 (right-handed molecule). The other (figure 5*d*), which has a period of about 8·3 A, is constructed of alternate units of types 1 and 2. Since neither of these periods agrees exactly with the known period of rubber, it does not appear possible to decide which model is correct; further, it appears that certain angles in the *cis* (rubber) chain are not quite the same as those in the *trans* (gutta-percha) chain. Nevertheless, the position is simplified by the above considerations to the extent that only two chain types need be considered.

Crystal structure determination

Rubber may be caused to crystallize either by stretching a specimen to several times its original length, whereby crystals with one axis parallel to the direction of stretching are produced, or (usually) by cooling below 0° C, when crystals with random orientation are formed. The correspondence between the X-ray diffraction patterns of stretched and frozen specimens shows that the crystal structure (that is, the molecular structure and the arrangement of the molecules) is the same in both types (Clark, Wolthuis & Smith 1937). On account of the complexity of the structure, only oriented specimens can be used for structure determination. X-ray diffraction patterns of stretched threads were first taken; these patterns correspond closely with those obtained by Morss (1938) and Sauter (1937). In addition, a very marked double orientation effect was obtained by stretching thin sheets (Gehman & Field 1939); diffraction patterns from these sheets provided valuable information on the orientation of the crystal planes responsible for the different spots, and thus made it possible to check the indices.

Several different unit cells have been suggested to account for the X-ray diffraction patterns of stretched rubber. Mark & Susich (1928) and Sauter (1937) suggested orthorhombic four-molecule (eight isoprene unit) cells of different dimensions, while Lotmar & Meyer (1936) proposed a monoclinic four-molecule cell. Morss (1938) concluded that all the diffraction spots, except one very weak one, fit a four-molecule orthorhombic cell with sides $a_0 = 12{\cdot}5$ A, $b_0 = 8{\cdot}9$ A and $c_0 = 8{\cdot}1$ A better than any other; this conclusion has been confirmed by the present work, and the intensities of spots from doubly oriented specimens confirm the indices assigned on this basis. Morss suggested that the additional very weak spot demands an eight-molecule cell which might be either orthorhombic or monoclinic; actually this is not the only way of accounting for the extra spot, as we shall see later. It was decided to put aside this question and accept provisionally Morss's four-molecule orthorhombic cell for preliminary work on the structure.



(a)

(b) (c)

FIGURE 6. Crystal structure of rubber. (*a*) Seen along *c* axis. (*b*) Seen along *b* axis. (*c*) Seen along *a* axis.

There are very few equatorial spots on the fibre diagrams, and, as Morss pointed out, the absences indicate that, seen along the fibre axis (the *c* axis), all the molecules look alike and form an arrangement which can be referred to a one-molecule projected pseudo-cell. This very much simplifies the problem. The two molecular models (or rather their end-on projections as shown at the bottom of figure 5*c*, *d*) were tried in turn in this one-molecule pseudo-cell; it was found that correct intensities for the diffractions from *hk*0 planes (the equatorial spots) cannot be obtained with model 5*d*, but that approximately correct intensities can be obtained with model 5*c*; the latter is accordingly indicated as the correct structure of rubber molecules. (For the arrangement of the molecules in the *ab* projection, see figure 6*a*.)

The *h*0*l* intensities were considered next; here, the absence of diffracted beams from planes with *h* odd indicates an apparent halving of the *a* axis for this viewpoint; thus, only two molecules have to be considered. The approximate atomic positions along *a* were already known. For reasons similar to those already given in the case of gutta-percha, there are likely to be equal numbers of 'up' and 'down' molecules in the crystal; hence, of the two molecules just mentioned, one is upside down with respect to the other (see figure 6*b*). Relative positions of the two molecules giving approximately correct *h*0*l* intensities were easily found.

It was then necessary to consider the space group and the question of the extra spot which does not fit the four-molecule orthorhombic cell. The extra spot could be accounted for either by a four-molecule monoclinic cell, or by an eight-molecule cell which might be either orthorhombic or monoclinic. The former possibility was naturally considered first. The extra spot, which is rather diffuse, is on the second layer line, and its spacing is nearest to that of $\bar{4}02$ and $\bar{3}22$ of the provisional cell; by making the angle between the *a* and *c* axes 92° instead of 90°, the extra spot can be indexed as $\bar{4}02 + \bar{3}22$ of a monoclinic cell; and the rest of the spots on the photographs also fit this cell within the limits of error of measurement. In this four-molecule monoclinic cell, the only likely arrangement, in view of the absent reflexions and the already known approximate molecular positions, is $P\,2_1/a$. Atomic positions in $P\,2_1/a$, giving approximately correct relative intensities for all the diffraction spots on the photographs, were finally found. It was found necessary to admit considerable distortions of the 'ideal' model (figure 5*c*) in order to obtain correct intensities.

Description of crystal structure

The unit cell is monoclinic with $a_0 = 12{\cdot}46$ A, $b_0 = 8{\cdot}89$ A, $c_0 = 8{\cdot}10$ A, $\beta = 92°$. Four long-chain molecules pass through this cell parallel to the *c* axis. The space group is $P\,2_1/a$; there are two left-handed molecules (one 'up' and one 'down') and two right-handed molecules (one 'up' and one 'down'). The co-ordinates *xyz* of the carbon atoms of one left-handed 'up' molecule are the following:

C_1	0·753*a*	C_2	0·854*a*	C_3	0·845*a*	C_4	0·745*a*	C_5	0·968*a*	*x*
(CH_2)	0·899*b*	(C)	0·865*b*	(CH)	0·905*b*	(CH_2)	0·959*b*	(CH_3)	0·876*b*	*y*
	0·802*c*		0·703*c*		0·542*c*		0·457*c*		0·773*c*	*z*
C_6	0·744*a*	C_7	0·644*a*	C_8	0·659*a*	C_9	0·757*a*	C_{10}	0·532*a*	*x*
(CH_2)	0·834*b*	(C)	0·874*b*	(CH)	0·905*b*	(CH_2)	0·834*b*	(CH_3)	0·828*b*	*y*
	0·326*c*		0·215*c*		0·052*c*		0·975*c*		0·268*c*	*z*

The atoms of the other three molecules are at the other equivalent positions of space group $P\,2_1/a$:

$$-x,\ -y,\ -z;\quad \tfrac{1}{2}-x,\ \tfrac{1}{2}+y,\ -z;\quad \tfrac{1}{2}+x,\ \tfrac{1}{2}-y;\quad z.$$

The complete structure is illustrated in figure 6*a*, *b* and *c*. Details of the structure of one molecule are shown in figure 7.

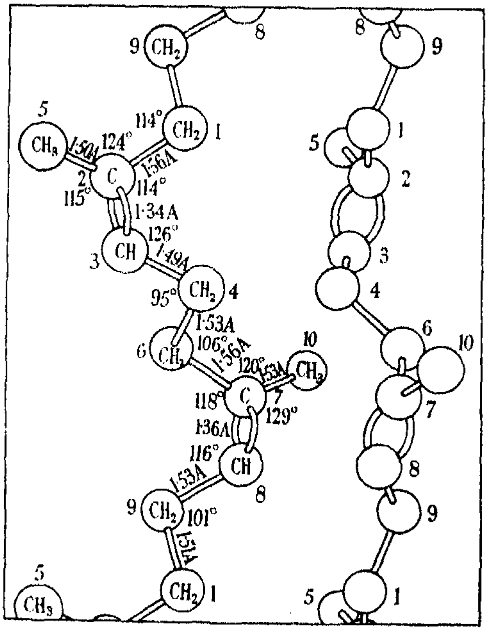


FIGURE 7. Left-handed rubber molecule, seen from two different viewpoints.

The atomic positions recorded above give approximately correct intensities for all the X-ray diffraction spots; but, in view of the complexity of the structure (in which there are 30 parameters), and the limited information obtainable from fibre photographs, the accuracy is naturally inferior to that obtainable in structure determinations based on single crystal photographs. For this reason, the atomic positions given here must be regarded as approximate only. The same applies to the other structures given in this work, but in a lesser degree on account of the comparative simplicity of the structures of β gutta-percha and polychloroprene, in which there are only 15 parameters.

The geometry of the rubber molecule, as it exists in the crystal (figure 7), shows considerable deviations from the 'ideal' model of figure 5*c*. In the first place, the two isoprene units which make up the identity period of an individual molecule are not identical in configuration. (This fact is revealed directly by the occurrence of a faint 001 diffraction spot, which was not seen by Morss (1938), but was definitely obtained by Sauter (1937) and in the present work.) Secondly, the double-bond chain units $C—C=C—C$ are not quite planar. Thirdly, the angles between the two $CH_2—CH_2$ chain bonds ($C_9—C_1$ and $C_4—C_6$) and the $CH_2—C=CH$ planes are different from each other and from the corresponding angles in β gutta-percha. Also the methyl groups lie considerably off the $CH_2—C=CH$ planes, though this feature was anticipated through knowledge of the β gutta-percha structure. The explanation of the last-mentioned feature is, no doubt, similar to that given in the case of β gutta-percha—it is due mainly to repulsion between the CH_3 groups and particular CH_2 chain groups. The other distortions appear to be due to the packing together of very awkwardly shaped molecules; each molecule is distorted from its 'ideal' form in order to pack better with its neighbours. Inspection of a model of the structure shows that all the distortions are reasonably explained (qualitatively) on this basis. This conclusion, that all the distortions which have to be admitted in order to account for the intensities of the X-ray diffraction spots are exactly those to be expected on account of packing difficulties, gives confidence that the whole structure is substantially correct, though individual atomic positions may be rather approximate.

There is one other X-ray diffraction effect of rubber crystals which should perhaps be mentioned. Diffraction spots at very small angles, indicating spacings of 54 and 108 A, have been reported (Clark, Wolthuis & Smith 1937; Clark, Gross & Smith 1939). These spacings are many times greater than the unit cell dimensions, and suggest some kind of superlattice. It seems possible that these spots, like the extra spots of aluminium, sodium chloride and magnesium oxide crystals investigated by Preston (1939), are due to thermal vibrations in the lattice. Atomic and molecular vibrations are very slow compared with the frequencies of X-rays; consequently, with regard to X-rays, the atoms in the lattice are in effect at rest but displaced from their normal positions, some being displaced in one direction and others in other directions, according to the wave pattern in the lattice. A super-lattice effect is thus produced by the thermal wave pattern in the lattice.

(c) POLYCHLOROPRENE

Chain form

It has already been remarked that the identity period of stretched polychloroprene (4·79 A) is almost identical with that of β gutta-percha (4·72 A), and therefore it probably has the same chain structure, that is to say, it has a *trans* double-bond configuration and the same single-bond configuration as β gutta-percha.

Crystal structure determination

Polychloroprene, like rubber, is usually non-crystalline at room temperature;* on stretching, oriented crystals are formed, and on release the molecules disarrange themselves again and the material again becomes non-crystalline. X-ray diffraction photographs of stretched specimens were taken under similar conditions to those of β gutta-percha and rubber. As compared with those of rubber, the photographs have a higher background intensity; also the spots are longer arcs, showing that the orientation of the crystals is less perfect.

All the diffraction spots fit a four-molecule orthorhombic cell whose a axis is longer and whose b axis is shorter than those of β gutta-percha. The $hk0$ intensities (those of the equatorial spots on the normal fibre photograph) were considered first; it was found possible to account for these only by the arrangement shown in figure 8*a*, which is analogous to that of β gutta-percha, though the orientation of the molecules is considerably different. In view of this arrangement, the space group of the whole structure is evidently $P\,2_1 2_1 2_1$, if we accept the arguments advanced in the case of β gutta-percha. The c ordinates of the atoms were found by consideration of the intensities of remaining spots; the data for this purpose are rather scanty, and therefore the positions of the atoms along the c axis are known with less accuracy than in the case of β gutta-percha.

Description of crystal structure

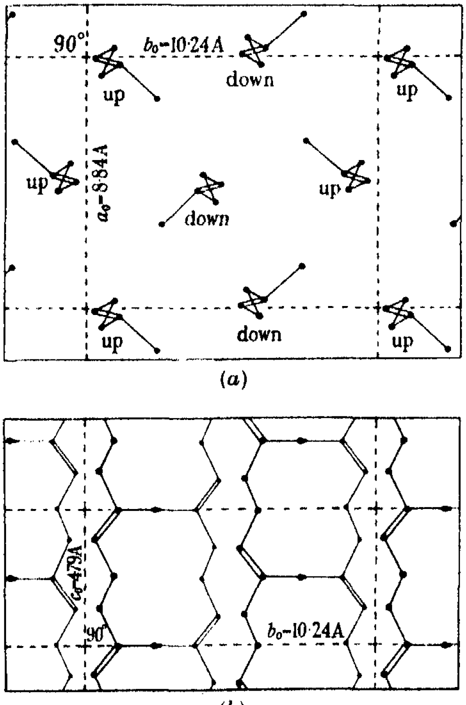
The unit cell is orthorhombic, with $a_0 = 8{\cdot}84$ A, $b_0 = 10{\cdot}24$ A, $c_0 = 4{\cdot}79$ A. Four long-chain molecules pass through this cell parallel to the c axis. The space group is $P\,2_12_12_1$ (the whole structure being completely analogous to that of β gutta-percha), and the co-ordinates *xyz* of the carbon and chlorine atoms of one right-handed 'up' molecule are the following:

---

* Some specimens are partly crystalline at room temperature. Rubber is also variable in this respect (Barnes 1937; Clark, Gross & Smith 1939).

C_1	$0{\cdot}077a$	C_2	$0{\cdot}033a$	C_3	$0{\cdot}008a$	C_4	$0{\cdot}970a$	Cl	$0{\cdot}167a$	x
(CH_2)	$0{\cdot}054b$	(C)	$0{\cdot}115b$	(CH)	$0{\cdot}034b$	(CH_2)	$0{\cdot}099b$		$0{\cdot}243b$	y
	$0{\cdot}278c$		$0{\cdot}000c$		$0{\cdot}787c$		$0{\cdot}509c$		$0{\cdot}000c$	z

The atoms of the other three molecules of the unit cell are at the other equivalent positions of $P\,2_12_12_1$ (see β gutta-percha structure). The complete structure is illustrated in figure 8 *a*, *b*; the similarity to that of β gutta-percha will be evident from these diagrams. All four molecules in the cell are identical—either all right-handed (in the example given here) or all left-handed, not mixed. Presumably there are equal proportions of left- and right-handed crystals in any specimen of stretched polychloroprene.



(*a*)

(*b*)

Figure 8. Crystal structure of polychloroprene, seen (*a*) along *c* axis, (*b*) along *a* axis.

All interatomic distances and bond angles are normal. The distance C—Cl is 1·77 A, which is a little greater than the distance in CCl_4 (1·75 A) and in chlorinated ethylenes (1·67–1·73) determined by electron diffraction in gases (Pauling & Brockway 1935; Brockway, Beach & Pauling 1935). The double-bond unit C—C=C—C is very nearly planar, and the CH_2—CH_2—C plane is at about 120° to the CH_2—C=CH plane (as compared with 115° in β gutta-percha). The C—Cl bond is not in the CH_2—C=CH plane, but is displaced by an amount (40°) even greater than the displacement of the CH_3 group in β gutta-percha (24°). This displacement is no doubt due, partly at any rate, to repulsion between the chlorine atom and chain CH_2 group C_4. Other influences, which may account for the larger angle in polychloroprene, are: distortion of the molecules by others packed round them (as in rubber) and the interaction of C—Cl dipoles in adjacent molecules.

The remarkable non-planar chain forms of the molecules studied in this paper will be discussed in relation to the stereochemistry of aliphatic molecules in general in Part II, and the bearing of the results on the question of molecular flexibility and the molecular basis of rubber-like properties in Part III of this work.

APPENDIX

Graphical method of determining intensities for hkl planes

Bragg & Lipson (1936) devised charts for determining the structure factors for planes of the type *hk*0, 0*kl* and *h*0*l*. These can also be used for shortening the work of calculating the structure factors for *hkl* planes.

For example, in β gutta-percha and polychloroprene the space group is $P\,2_1 2_1 2_1$, for which the structure factor is $\sqrt{(A^2+B^2)}$, where

$$A = \Sigma 4f \cos 2\pi\left(hx-\frac{h-k}{4}\right)\cos 2\pi\left(ky-\frac{k-l}{4}\right)\cos 2\pi\left(lz-\frac{l-h}{4}\right)$$

(summed for all atoms),

$$B = -\Sigma 4f \sin 2\pi\left(hx-\frac{h-k}{4}\right)\sin 2\pi\left(ky-\frac{k-l}{4}\right)\sin 2\pi\left(lz-\frac{l-h}{4}\right)$$

(summed for all atoms).

The xy positions of the atoms are plotted on a transparent square chart in the manner given by Bragg & Lipson; the z positions are plotted in a separate strip. To find A, the product of the first two cosine terms is read off on a chart giving $\cos 2\pi hx \cos 2\pi ky$ (plane group Pba) by suitably displacing the origin; for instance, for plane 231, $2\pi\left(\frac{h-k}{4}\right)$ is $-90°$ and $2\pi\left(\frac{k-l}{4}\right)$ is $+180°$; using the 230 chart, the origin of the atomic position chart is shifted 90° backwards along the x axis and 180° forwards along the y axis. The last cosine term is read off on the strip, using a one-dimensional chart giving $\cos 2\pi lz$; for plane 231, $2\pi\left(\frac{l-h}{4}\right)$ is $-90°$, and the origin of the strip must therefore be displaced backwards 90°. The only calculation is then the multiplication of the two graphically estimated figures and f, the diffracting power of the atom. Since terms are obtained in a similar way from the same charts.

This procedure is of course longer than for planes of type $hk0$, etc., but it is very much shorter than straightforward calculation of hkl structure factors.

The space group of rubber is $P\,2_1/a$, for which the structure factor is

$$A = 4\cos 2\pi\left(hx + lz + \frac{h+k}{4}\right)\cos 2\pi\left(ky - \frac{h+k}{4}\right) \quad (B = 0).$$

The x and z positions of the atoms are plotted on square chart, and the first cosine term read off on the Bragg and Lipson chart for plane group $P\,2$, the origin being displaced by $2\pi\left(\frac{h+k}{4}\right)$; the y positions are plotted on a strip as in the previous example, and the second term read off, again after the appropriate displacement of the origin; the two terms are then multiplied.

REFERENCES

Barnes 1937 *Canad. J. Res.* **15**, 156.
Bragg & Lipson 1936 *Z. Kristallogr.* **95**, 323.
Brockway, Beach & Pauling 1935 *J. Amer. Chem. Soc.* **57**, 2693.
Bunn 1939 *Trans. Faraday Soc.* **35**, 482.
Carothers, Williams, Collins & Kirby 1931 *J. Amer. Chem. Soc.* **53**, 4203.
Clark, Gross & Smith 1939 *Bur. Stand. J. Res.* **23**, 1.
Clark, Wolthuis & Smith 1937 *Bur. Stand. J. Res.* **19**, 497.
Fuller 1936 *Industr. Eng., Chicago*, **28**, 907.
Gehman & Field 1939 *J. Appl. Phys.* **10**, 564.
Guth 1937 *Z. Elektroch.* **43**, 683.
Harries 1905 *Ber. dtsch. Chem. Ges.* **38**, 3985.
Harries 1912 *Z. angew. Chem.* **25**, 1457.
Hauser & Susich 1931 *Kautschuk*, **7**, 120, 125, 145.
Hengstenberg 1928 *Z. Kristallogr.* **67**, 583.
Katz 1925 *Naturwissenschaften*, **13**, 411.
Kuhn 1936 *Kolloidzschr.* **76**, 258.
Kuhn 1939 *Kolloidzschr.* **87**, 3.
Lotmar & Meyer 1936 *Mh. Chem.* **69**, 115.
Mack 1934 *J. Amer. Chem. Soc.* **56**, 2757.
Mark 1939 *Chem. Rev.* **25**, 121.
Mark & Susich 1928 *Kolloidzschr.* **46**, 11.

Meyer & Mark 1928 *Ber. dtsch. Chem. Ges.* **61**, 1939.
Morss 1938 *J. Amer. Chem. Soc.* **60**, 237.
Müller 1928 *Proc. Roy. Soc.* A, **120**, 437.
Pauling & Brockway 1935 *J. Amer. Chem. Soc.* **57**, 2684.
Preston 1939 *Proc. Roy. Soc.* A, **172**, 116.
Robertson 1938 *J. Chem. Soc.* p. 131.
Sauter 1937 *Z. phys. Chem.* B, **36**, 405.
Staudinger 1930 *Ber. dtsch. Chem. Ges.* **63**, 734, 921.
Storks 1938 *J. Amer. Chem. Soc.* **60**, 1753.
Thompson 1939 *Trans. Faraday Soc.* **35**, 697.

TABLES OF RESULTS

In accordance with a recent recommendation of the Royal Society (30 March 1939), tables of figures in connexion with the crystal structure determinations are not included in the paper itself, but are given separately.

The figures given in the following tables are the relative intensities calculated for X-ray diffractions from the various crystal planes, assuming the atomic positions given in the paper. The figures are proportional to $p[\theta]F^2$, where F is the structure factor (see below), $[\theta]$ the angle factor (for instance, $\frac{1+\cos^2\theta}{\sin^2\theta}$ for the equatorial spots of a fibre photograph), and p the number of co-operating crystal planes. For non-equatorial spots Cox and Shaw's correction factors were used (*Proc. Roy. Soc.* A, 1930, **127**, 71).

The structure factor for space group $P\,2_1 2_1 2_1$ (the space group of β gutta-percha and polychloroprene) is $\sqrt{(A^2+B^2)}$, where

$$A = \Sigma 4f \cos 2\pi\left(hx - \frac{h-k}{4}\right)\cos 2\pi\left(ky - \frac{k-l}{4}\right)\cos 2\pi\left(lz - \frac{l-h}{4}\right),$$

$$B = -\Sigma 4f \sin 2\pi\left(hx - \frac{h-k}{4}\right)\sin 2\pi\left(ky - \frac{k-l}{4}\right)\sin 2\pi\left(lz - \frac{l-h}{4}\right).$$

For space group $P\,2_1/a$ (that of rubber), the structure factor is

$$A = \Sigma 4f \cos 2\pi\left(hx + lz + \frac{h+k}{4}\right)\cos 2\pi\left(ky - \frac{h+k}{4}\right), \quad B = 0.$$

For f, the diffracting power of an atom, the values given for uncharged carbon and chlorine atoms by Bragg (*The crystalline state*, **1**) were used. The diffracting powers of CH, CH_2 and CH_3 groups were assumed to have diffracting powers $\frac{7}{6}$, $\frac{8}{6}$, and $\frac{9}{6}$ that of carbon itself, these figures being proportional to the number of electrons in each group. This is strictly true only for small angles of diffraction, but at any angle it is a better approximation to the truth than the assumption of equal diffracting powers for these groups.

No corrections have been made for the thermal vibrations of the atoms.

1. β GUTTA-PERCHA

The three columns in each section give: indices, calculated intensity, observed intensity.

A. *Normal fibre photograph*

	*hk*0	
110	0·2	—
020	4·6	vw
120	325·0	vs +
200	207·0	vs
210	3·9	—
130	0·3	—
220	5·5	vw
040	24·8	m
230	1·7	ms
140	38·1	
310	0·1	—
320	9·8	m
240	12·2	
150	0·2	—
330	0·6	—
250	0·5	—
060	7·2	w
340	6·0	w
400	7·0	
410	2·6	m
160	28·6	
420	0·6	—
260	6·9	vvw
350	0·7	
430	1·6	
170	0·1	—
440	0·2	—
360	5·2	vvw
270	0·1	—
510	0·1	
520	0·1	—

	*hk*1	
011	74·0	ms
101	0·6	—
111	8·9	vw
021	0·2	—
121	10·7	—
031	15·9	w
201	5·6	
211	32·8	ms
131	3·1	—
221	1·2	—
041	1·8	—
231	6·7	vw
141	1·4	
301	1·0	w
311	7·1	
321	2·0	—
051	1·0	—
241	3·1	—
151	0·2	—
331	3·6	—
251	3·5	—
061	2·2	—
341	1·0	—
401	2·3	—
411	13·6	w
161	1·8	
421	1·6	—

	*hk*2	
112	26·0	w
022	52·0	
122	2·0	—
032	0·2	w
202	11·2	
212	6·0	
132	9·7	
222	29·1	
042	0·0	—
232	8·7	—
142	2·5	—

B. *Fibre axis oscillating* 90°–45° *to X-ray beam*

	hkl	
002	5·4	mw
012	1·8	—
102	0·8	
013	1·5	vvw
103	0·2	
113	0·5	
023	0·0	—
004	0·5	vw
014	3·1	
104	0·0	

For all réflexions *h*00 with *h* odd, 0*k*0 with *k* odd, and 00*l* with *l* odd, the observed and calculated intensities are nil.

vs + = extremely strong, vs = very strong, s = strong, ms = medium strong, m = medium, mw = medium weak, vw = very weak, vvw = extremely weak.

2. RUBBER

A. *Normal fibre photograph*

	*hk*0	
110	1·5	—
200	64·3	s
210	0·3	—
020	0·2	—
120	254·0	vs +
310	0·6	—
220	5·0	—
400	15·0	m
320	4·0	
410	0·2	—
130	0·1	—
230	0·2	—
420	0·2	—
330	0·4	—
510	0·0	
040	25·7	m
140	0·0	—
520	4·6	
430	0·0	—
240	1·2	—
600	0·0	
610	0·0	—
340	0·0	—
530	0·0	—
620	0·0	—
440	0·0	—
150	1·4	—
710	0·0	—
250	0·0	—
630	0·0	
540	1·8	—
720	0·0	
350	0·8	—
800	4·6	vvw
450	0·0	
730	0·0	
810	0·0	—
640	0·4	—
060	0·1	—
160	1·8	—
820	0·0	—
550	0·0	—
260	0·1	—
360	0·3	—
740	0·0	—
900	0·0	—
830	0·0	—
910	0·0	—
650	0·0	—
460	2·0	—
920	15·0	vw
840	1·0	—
560	0·3	—
170	0·0	—
750	0·0	—
930	0·0	—
10.0.0	7·3	vvw
270	0·0	—
10.1.0	0·0	—
370	0·0	—
660	0·6	—

	*hk*1	
011	0·7	—
$\bar{1}11$	2·0	vw
111	2·3	
$\bar{2}01$	22·0	s
201	26·4	
$\bar{2}11$	3·3	—
211	1·6	
021	2·9	—
$\bar{1}21$	31·1	s
121	22·2	
311	6·5	mw
311	4·1	
221	1·8	—
221	1·5	
$\bar{4}01$	0·0	—
401	0·0	vw
$\bar{3}21$	0·5	
321	0·0	
$\bar{4}11$	0·2	
031	2·8	
$\bar{1}31$	0·2	—
411	0·5	
131	0·3	
$\bar{2}31$	2·4	mw
231	2·8	
$\bar{4}21$	0·0	—
421	1·0	
$\bar{3}31$	0·4	vvw
$\bar{5}11$	0·0	
331	1·0	
511	0·3	
041	0·7	—
$\bar{5}21$	1·0	—
$\bar{1}41$	2·1	
141	0·0	
$\bar{4}31$	4·0	vw
521	0·4	
431	0·0	
$\bar{6}01$	0·4	w
$\bar{2}41$	5·8	
241	1·0	
601	0·6	
$\bar{6}11$	0·5	—
611	0·4	
$\bar{3}41$	0·0	—
341	1·5	
531	0·3	—
$\bar{6}21$	0·0	—
531	0·0	
621	0·9	—
$\bar{4}41$	0·4	—
441	0·6	—
051	0·0	—
$\bar{1}51$	2·0	vw
151	0·1	
$\bar{7}11$	1·3	
711	0·9	
$\bar{6}31$	0·0	—
$\bar{2}51$	0·0	
251	0·0	
631	0·4	—
$\bar{5}41$	0·6	vw
721	0·4	
541	0·0	
$\bar{3}51$	1·5	
721	0·6	
351	2·8	
$\bar{8}01$	1·0	vw
731	0·2	
$\bar{4}51$	0·0	
801	0·9	
451	0·0	
731	0·2	

2. RUBBER (*continued*)

A. *Normal fibre photograph* (*continued*)

	*hk*2	
012	6·9	mw
$\bar{1}12$	4·3	
112	1·4	—
$\bar{2}02$	23·6	ms
202	29·4	
$\bar{2}12$	0·4	—
212	2·9	—
022	5·8	m
$\bar{1}22$	0·0	
122	0·6	—
$\bar{3}12$	2·8	m
$\bar{2}22$	0·0	
312	4·6	
222	0·0	
$\bar{4}02$	1·1	vvw
$\bar{3}22$	1·3	
402	0·6	—
$\bar{4}12$	6·7	mw
032	0·1	
322	2·0	
$\bar{1}32$	0·2	
132	0·8	
412	6·5	
$\bar{2}32$	0·2	vvw
232	0·0	
$\bar{4}22$	0·0	
422	0·0	—
$\bar{5}12$	8·2	m
$\bar{3}32$	0·8	
332	0·8	
512	4·1	
042	1·7	m
$\bar{5}22$	1·6	
$\bar{1}42$	0·7	
$\bar{4}32$	1·0	
142	0·0	
522	2·3	
432	3·0	
$\bar{6}02$	0·0	
$\bar{2}42$	0·0	—
242	0·1	
$\bar{6}12$	1·6	vvw
602	0·0	
612	2·8	

	*hk*3	
013	0·0	—
$\bar{1}13$	2·0	mw
$\bar{2}03$	1·0	
113	20·0	
203	1·5	—
$\bar{2}13$	0·0	
213	0·0	—
023	6·3	mw
$\bar{1}23$	3·7	
123	8·8	
$\bar{3}13$	1·0	
$\bar{2}23$	0·4	—
313	0·2	—
223	7·2	
$\bar{4}03$	2·6	
$\bar{3}23$	4·7	mw
$\bar{4}13$	0·0	
403	2·0	
033	2·7	
323	1·5	
$\bar{1}33$	1·6	
133	0·7	
413	0·0	
$\bar{2}33$	0·0	—
233	0·0	—
$\bar{4}23$	1·1	

	*hk*4	
014	5·3	w
$\bar{1}14$	0·0	
$\bar{2}04$	0·0	
114	0·0	
204	0·5	—
$\bar{2}14$	0·0	
214	0·0	—
024	0·0	—
$\bar{1}24$	5·0	mw
124	6·3	
$\bar{3}14$	0·0	
$\bar{2}24$	0·3	
314	0·0	—
224	1·4	
$\bar{4}04$	0·8	—
$\bar{3}24$	0·0	vw
$\bar{4}14$	1·3	
034	0·9	
404	0·4	
$\bar{1}34$	0·0	
324	0·6	
134	0·8	
414	0·4	

B. *Fibre axis oscillating* 90°–45° *to X-ray beam*

	00l	
001	0·2	vvw
002	8·0	mw
003	0·0	—
004	9·2	mw

	hk5	
005	0·0	—
015	0·1	—
$\bar{1}15$	0·0	vw
205	1·1	
115	0·0	
$\bar{2}15$	0·0	—
205	0·7	
215	0·0	—
025	0·0	vw
$\bar{1}25$	2·2	
$\bar{3}15$	0·3	
125	0·8	
$\bar{2}25$	0·1	
315	0·6	
225	0·0	

For all reflexions $0k0$ with k odd, $h00$ and $h0l$ with h odd, the observed and calculated intensities are nil.

3. POLYCHLOROPRENE

A. *Normal fibre photograph*

	hk0	
110	1·8	w
020	0·5	—
200	5·5	vs
120	25·8	
210	18·6	s
220	15·7	s
130	1·2	—
310	1·3	—
230	3·2	w
040	1·8	w
320	2·2	
140	3·5	vw
330	0·2	—
240	3·1	vw
400	0·0	
410	2·9	vvw
420	2·5	vvw
150	0·1	—
340	1·5	—
430	4·8	mw
250	4·0	
510	0·1	—
060	2·9	vvw
350	0·0	vvw
160	0·2	
440	1·3	
520	0·2	
260	0·0	—
530	0·0	—

	hk1	
011	32·0	ms
101	0·0	—
111	0·2	—
021	0·2	—
201	4·4	w
121	0·7	
211	4·0	
031	7·0	mw
221	7·3	
131	0·2	—
301	0·1	—
311	0·1	
231	0·3	—
041	0·0	—
321	0·9	
141	0·0	—

	hk2	
112	1·0	mw
022	23·0	
202	1·3	—
122	1·6	
212	8·0	
032	2·2	—
222	0·6	

B. *Fibre axis oscillating* 90°–45° *to beam*

	hk2	
002	3·3	mw
012	0·1	—
102	0·0	—

For all reflexions $h00$ with h odd, $0k0$ with k odd, and $00l$ with l odd, the observed and calculated intensities are nil.

Molecular structure and rubber-like elasticity

II. The stereochemistry of chain polymers

By C. W. Bunn, B.A., B.Sc.

The Research Department of Imperial Chemical Industries Ltd., Northwich

(*Communicated by Sir William Bragg, F.R.S.—Received* 19 *May* 1941)

A survey of the known structures of molecules containing two or more singly linked carbon atoms has shown that the staggered bond configuration is favoured. It is suggested that these molecules may be examples of a general rule—the 'principle of staggered bonds'—which specifies the bond positions for successions of singly linked carbon atoms, just as the 'tetrahedral principle' specifies the bond positions on any one carbon atom.

If this is true, it is possible to predict molecular configurations. The number of alternative configurations, all obeying the principle of staggered bonds, increases rapidly with the number of carbon atoms, but the most probable configuration of a particular molecule can sometimes be predicted on geometrical grounds.

The simpler types of carbon chain are described. The chain-type in a crystalline long-chain polymer may sometimes be deduced from a knowledge of the identity period of the molecules, which is readily found from X-ray or electron diffraction patterns.

The principle of staggered bonds is also obeyed in molecules containing singly linked nitrogen, oxygen and sulphur atoms. The configurations of hetero-atomic chains containing these atoms are therefore of the same types as carbon chains.

In molecules containing double bonds, the principle of staggered bonds may still be operative, a double bond being regarded in the familiar simple geometrical way as two distorted single bonds. The non-planar chain forms of gutta-percha, rubber and polychloroprene are attributed to the co-operation of two influences: the principle of staggered bonds, and the repulsion between methyl or chlorine side substituents and particular CH_2 groups in the chains.

In Part I of this work (Bunn 1941) the determination of the crystal structures of three long-chain polymers by interpretation of X-ray diffraction photographs was described. In all three crystals—β gutta-percha, rubber and polychloroprene—the molecules have non-planar zigzag chain forms and are asymmetric. It is now necessary to consider the bearing of the new knowledge of molecular geometry on the possibility of understanding rubber-like properties in terms of molecular physics. It is widely believed that the flexibility, softness and other characteristic properties of rubber-like substances are in some way due to the flexibility of the molecules themselves.

Long-chain molecules owe their potential flexibility to the swivelling of the chain units round the single bonds as axes, and it is therefore necessary to consider which bond positions are the most stable and what hindrances there are to rotation away from these positions. The present paper deals chiefly with the question of the most stable bond positions. The enquiry has interest, not only in relation to the problem of the origin of rubber-like properties, but also because it opens the way to a systematic consideration of chain types. There is already evidence that in many crystalline long-chain polymers, such as rubber hydrochloride (Gehman, Field & Dinsmore 1938), polyisobutylene (Fuller, Frosch & Pape 1940) and some of the polyesters (Fuller 1940), the chains have not the fully extended plane zigzag form of polyethylene (Bunn 1939) but somewhat shortened (necessarily non-planar) forms. It should be possible to discover what these forms are by interpretation of X-ray diffraction photographs, but the difficulties are in some cases formidable; some assistance in the form of guiding principles for the construction of possible chain types is desirable. It is the purpose of this paper to show that there already exists sufficient evidence to suggest a general principle regulating bond positions in aliphatic molecules containing sequences of singly linked atoms. It will be called the principle of staggered bonds.

In the three molecules whose structures were determined in Part I, every fourth chain bond is a double bond; the question of bond positions is less simple for such molecules than it is for those in which all the bonds are single. The latter will therefore be considered first.

BOND POSITIONS IN SATURATED MOLECULES

A survey of all the well-established structures of simple, saturated organic molecules, whether in the gaseous state or in crystals, has been made, and has shown that the favoured positions of the bonds of two singly linked carbon atoms are as shown in figure 1. The tetrahedral angles between the bonds of any one carbon atom are preserved (within a few degrees), and the two trios of bonds *abc* and *a'b'c'* of two singly linked carbon atoms are staggered with respect to each other. Thus, for ethane, both the electron diffraction pattern (Pauling & Brockway 1937) and the infra-red absorption spectrum (Bartholomé & Karweil 1938) suggests that the two trios of hydrogen atoms spend most of their time in the staggered position; they undoubtedly rotate round the C—C bond as axis, but the rotation is not a revolution at constant speed, or a random rotation, but a rapid switching from one to another of the three equivalent positions.

In 1, 2-dichlorethane, $CH_2Cl.CH_2Cl$, both X-ray diffraction patterns (Ehrhardt 1932) and electron diffraction patterns (Beach & Palmer 1938) indicate that the halogen atoms spend most of their time at the maximum possible distance from each other, that is, with the C—Cl bonds at positions a, a' of figure 1. The same is true for the corresponding di-bromo-ethane, $CH_2Br.CH_2Br$, and for the chlorobromo compound, $CH_2Cl.CH_2Br$ (Beach & Turkevitch 1939). There is some indication that there may be alternative positions with the C—C—X planes at 80° to each other, in addition to the most frequent position with these planes at 180°. An angle of 80° means 20° deviation from alternative staggered positions ab' or ac' of figure 1; the deviation may be due to mutual repulsion of the halogen atoms. (The electric dipole moment of C—Cl is high.) That there are alternative positions in $CH_2Cl.CH_2Cl$ is also shown by the appearance in the Raman spectrum of a doublet for the C—Cl vibration; the two components of the doublet have different intensities at room temperature, showing that one position (the 180° position according to X-ray and electron diffraction patterns) is favoured, but the intensities become more nearly equal with rise of temperature, showing that the second position is visited relatively more frequently at higher temperatures (Kohlrausch 1932).

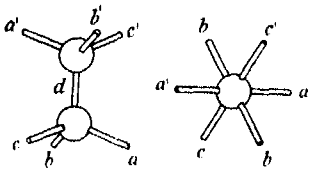
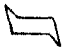
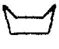


FIGURE 1. Bond positions in saturated molecules.

In 1, 1, 2-trichlorethane, $CHCl_2.CH_2Cl$, the angle between the planes Cl—C^1—C^2 and C^1—C^2—Cl is 70°—within 10° of the precise staggered position (Turkevitch & Beach 1939); and in $Br(CH_3)CH—CH(CH_3)Br$ the staggered position is again the most stable, the bromine atoms being in positions aa' of figure 1 (Stevenson & Schomaker 1939). In cyclohexane there are only two structures which allow tetrahedral angles between the carbon bonds; the 'chair'  and 'boat'  forms of Sachse. It is interesting to find that in gaseous cyclohexane the molecules have the 'chair' form, in which all the bonds are staggered (figure 2c); the 'boat' form, in which the bonds are opposed, does not occur (Pauling & Brockway 1937). In trimethyl methane (Beach & Stevenson 1938) and in tetramethyl methane (Pauling & Brockway 1937), there is some indication that the hydrogen atoms of the methyl groups may be most frequently at positions 30° from the staggered position, that is halfway between the staggered and opposed configurations. The evidence for all these molecules is provided by the electron diffraction patterns of the gases.

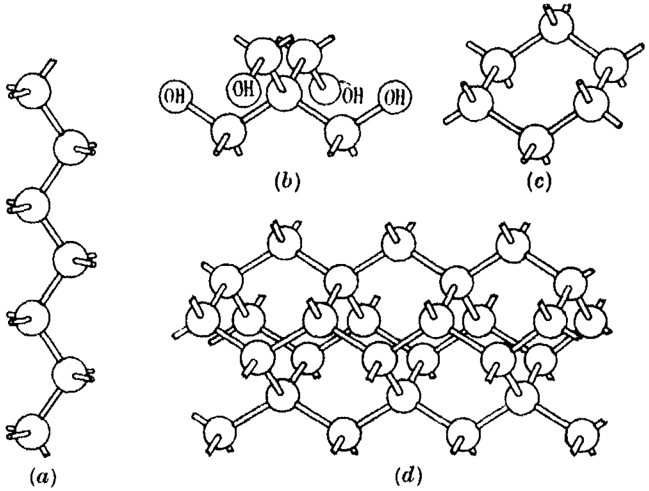


FIGURE 2. Structures of some singly-linked molecules. (*a*) Normal paraffin hydrocarbons. (*b*) Penta-erythritol. (*c*) Cyclo-hexane. (*d*) Diamond.

In crystals, likewise, the staggered configuration is favoured. The plane zigzag of carbon atoms in long-chain normal paraffin hydrocarbons such as poly-ethylene (Bunn 1939) is simply a repetition of the plane zigzag *ada′* (or *bdb′* or *cdc′*) of figure 1, and assuming that the hydrogen atoms are in the tetrahedral positions, all the bond orientations in the molecule (see figure 2*a*) conform to the model of figure 1. In penta-erythritol, $C(CH_2OH)_4$ (Llewellyn, Cox & Goodwin 1937), and its tetra-acetate (Goodwin & Hardy 1938) we find the same scheme of bond orientation (figure 2*b*), while in diamond (Bragg 1913) the scheme attains its most extended expression, a perfect diamond crystal (figure 2*d*) being simply a gigantic molecule composed of carbon atoms all linked together in precisely the manner shown in figure 1. The carbon atoms of succinic acid, $HOOC.CH_2.CH_2.COOH$, form a plane zigzag as in the normal paraffin hydrocarbons (Verweel & MacGillavry 1938). The same is true for the tartrate ion in Rochelle salt (Beevers & Hughes 1941).

In 1, 2-di-iodo-ethane, $CH_2I.CH_2I$, the distance between the iodine atoms (Klug 1935) leaves little doubt that they are at the maximum possible distance from each other, that is, the C—I bonds are at positions *aa′*

of figure 1. Similarly, the positions of the halogen atoms in the crystalline hexachloro and hexabromo derivatives of cyclohexane (Dickinson & Bilicke 1928) indicate strongly that the carbon ring (like that of gaseous cyclohexane) has the 'chair' form, in which the bonds of linked carbon atoms are staggered (figure 2*c*).*

Reason for the staggered configuration

The number of instances in which the staggered configuration is the most stable is sufficient to suggest that it may be a quite general rule which regulates the bond positions on successions of singly linked carbon atoms, just as Van 't Hoff's tetrahedral model specifies the bond positions on any one carbon atom. In this connexion it is relevant to enquire into the reason for the staggered configuration. The tetrahedral disposition of the bonds of any one carbon atom is primarily due, not to mutual repulsion of the atoms or groups held by the bonds, but to a strong orienting tendency of the bonds themselves (the shared electron links being most stable when tetrahedrally disposed), and it may be that the staggered configuration of the bonds of linked carbon atoms is due to a similar orienting tendency of the bonds themselves. Although there is no clear evidence on this point, one influence —the effect of electric dipole moments—may be discounted; in the molecule of 1, 1, 2-trichlorethane, $CHCl_2 . CH_2Cl$, all the electric dipole moments $H^+—C^-$ and $C^+—Cl^-$ favour the opposed configuration; that is to say, if the configuration were determined mainly by the dipole moments, the angle between the planes $Cl—C^1—C^2$ and $C^1—C^2—Cl$ would be 120°, giving the maximum distances between the chlorine atoms. Actually (figure 3) the angle is 70°, within 10° of the staggered bond position; there are two 70°

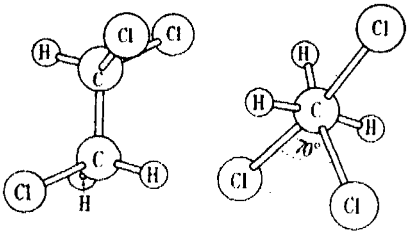


FIGURE 3. Molecular structure of 1, 1, 2 trichlorethane $CHCl_2 . CH_2Cl$.

* Since this paper was first written, Schomaker & Stevenson (*J. Chem. Phys.* 1940, 8, 637) have also pointed out that the staggered configuration is favoured, citing some of the evidence given here.

positions, between which oscillation occurs. Evidently the electric dipole effect is small; *the staggered bond configuration is therefore due either to an orienting tendency of the bonds themselves or to atomic interactions not attributable to electric dipole effects, or to both these factors.* Kistiakowsky, Lacher & Ransom (1938) have suggested that, quite apart from the question as to which configuration is the most stable, the potential barrier to rotation is due to interaction between the electron pairs constituting the bonds. It should perhaps be mentioned that quantum-mechanical calculations of the forces in the ethane molecule have recently been made by several investigators. Some of these appear to point to the opposed configuration as the more stable; but the problem is a complex one, and different authors do not agree about the contributions of various factors (see Gorin, Walter & Eyring 1939; Eucken & Schäffer 1939). The theoretical treatment must therefore be regarded as tentative. Meanwhile, the experimental evidence must be accepted, which is strongly in favour of the general stability of the staggered configuration. The barrier to rotation is 3000 cal. in ethane (Pitzer 1937) and at least 5000 cal. in 1, 1-dichlorethane and similar molecules (Beach & Palmer 1938; Beach & Turkevitch 1939).

Deviations from the staggered positions must be expected when the bonds hold atoms or groups of very different sizes, or still more when the atoms take part in certain ring structures. Thus, in cyclopentane, the carbon atoms form a plane ring, in which the bonds are opposed, not staggered (Pauling & Brockway 1937). The reason is, no doubt, that the preservation of the tetrahedral angles between the bonds on any one carbon atom is a more urgent necessity than the maintenance of the staggered bond scheme for successive carbon atoms.

PREDICTION OF STRUCTURES OF SATURATED MOLECULES

If the principle of staggered bonds is of general validity (with the reservations mentioned in the last paragraph), it can be used for predicting the possible configurations of saturated molecules. It will not, in general, indicate only one possible structure for any particular molecule, because alternative bond positions will (except in the simplest molecules) give rise to different configurations; but it will usually indicate a limited number of possible structures, each of which can be tried out in turn. A useful way of regarding the possible single-bonded structures comes from the realization that all single-bonded carbon structures which conform to figure 1 are portions of the diamond structure (figure 2*d*). It is not to be expected that the precise geometry of any particular structure can be predicted, because the interaction of different groups will give rise to distortions of the 'ideal' model; but an approximate prediction is better than no prediction at all.

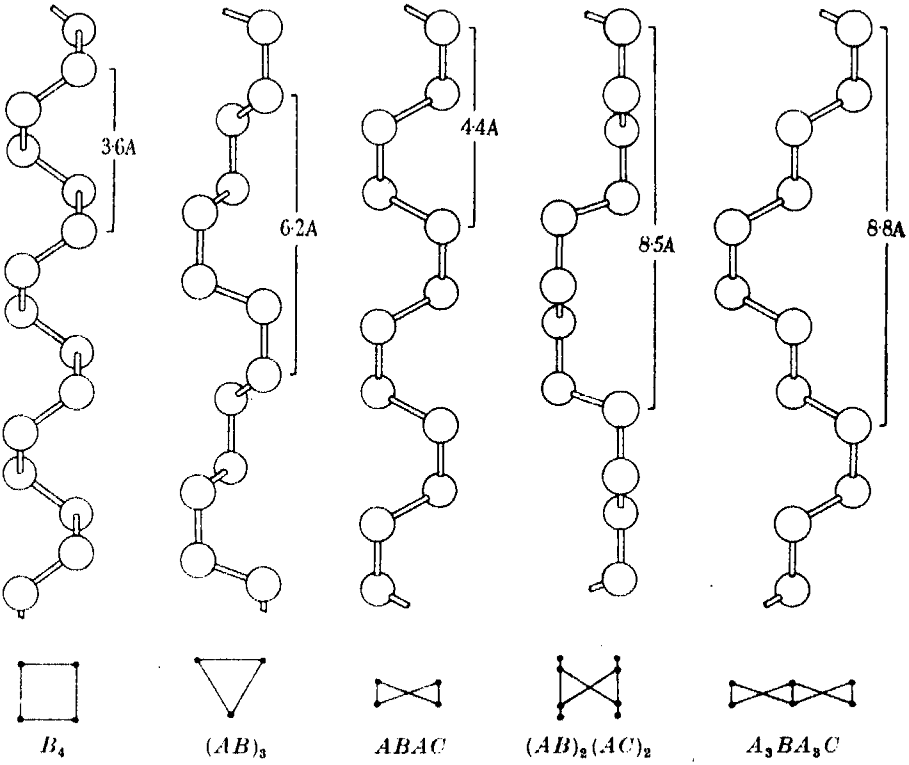
In attempting to predict the structure of any particular molecule, it will be necessary to consider first the possible alternative bond-sequences, to select if possible the most probable of these, and then to consider what distortions of the ideal model are likely to occur.

For large molecules there may be many possible bond sequences, but it will often happen that consideration of atomic radii and the clearances between different parts of the molecule will show that one bond sequence is more probable than all others. As an example, consider how the simple plane zigzag chain of the normal paraffin hydrocarbons arises. Any two chain bonds such as a' and d in figure 1 are at the tetrahedral angle to each other. For the next chain bond there are three possible positions, a, b and c. If it took up position b, the carbon atoms on a' and b would be only 2·9 A apart; since the radius of a CH_2 group is 1·8–2·05 A, it is evident that there would be considerable repulsion between the groups in these positions. The same is true for bond position c, which is exactly equivalent to b in this respect. In position a, however, the carbon atoms on a' and a are at the maximum possible distance from each other (3·9 A); this position is therefore more probable than the other two. The same applies to all the chain bonds, hence the most probable chain configuration is the plane zigzag (figure 2a). This is the structure which chains of CH_2 groups are likely to have at low temperatures, and thus it is not surprising that this is the structure we find in crystalline normal paraffin hydrocarbons.

NOMENCLATURE OF CHAIN TYPES

When side substituents are present on a carbon chain, the plane zigzag may not be the most probable form of chain. It will be useful to have a simple way of referring to different chain types. For any three chain bonds, we have seen that there are three different configurations, $a'da$, $a'db$ and $a'dc$; these will be called A, B and C respectively. Repetition of A gives the plane zigzag (figure 2a) which has an identity period of 2·53 A; this chain is best symbolized AA or A_2, since this indicates the two-atom identity period as well as the bond sequence. Other chain types are illustrated in Fig. 4. Repetition of B gives rise to a right-handed spiral chain $BBBB$ or B_4 with a four-atom period of 3·6 A; in the end-view of this chain, the atoms fall at the corners of a square. Repetition of C gives the corresponding left-handed spiral C_4.

Alternation of A and B gives rise to a left-handed spiral chain $(AB)_3$ with a six-atom period of 6·2 A; in the end-view, the chain atoms fall at the corners of an equilateral triangle. Alternation of A and C gives the corresponding right-handed spiral $(AC)_3$. Alternation of B and C gives the cyclohexane ring $(BC)_3$, already illustrated in figure 2c. Other simple chain types which are not spirals, but merely non-planar zigzags, are also shown in figure 4.



B_4 $(AB)_3$ $ABAC$ $(AB)_2(AC)_2$ A_3BA_3C

FIGURE 4. Some of the simpler single-bonded chain-types.

DETERMINATION OF CHAIN TYPE IN CRYSTALLINE POLYMERS

These chain types arising from regular repetitions of particular bond sequences are those likely to be found in crystalline chain polymers. Each chain type has its own characteristic identity period, and the identity periods of the simpler types are quite different from each other (see figure 4). Evidently it should be possible to find the type of chain in a polymer molecule by mere determination of the identity period of the molecule; this is quite easily found from the X-ray (or electron) diffraction patterns of crystalline chain polymers, since in a stretched or rolled out specimen the identity period of the crystals in the direction of stretching or rolling (calculated directly from the distance between the layer lines on the diffraction pattern) is the identity period of the molecules themselves. The figures given are 'ideal' identity periods, based on the precise staggered bond model of figure 1. Some departure from these figures may be expected, though not enough to lead to uncertainty for the simpler chain types.

When a chain polymer is found to have a long identity period, more than one chain type may appear eligible; for with increase in the number of atoms per period the number of different bond sequences (some with similar identity periods) increases rapidly. In such cases, knowledge of the chemical structure of the chain molecules may help towards the selection of the correct chain type; each chemical unit in the chain would be expected to have a similar configuration, and thus the chemical structure of the chain would be reflected in the bond sequence. (For this purpose, enantiomorphic configurations such as AB and AC are equivalent.) As an example, rubber hydrochloride, which is believed to be chemically

$$(-CH_2-CH_2-CH_2-C(CH_3)Cl-)_n,$$

is a crystalline substance with a molecular identity period of 9·1 A (Gehman, Field & Dinsmore 1938). The chain structure A_3BA_3C (figure 4) is suggested for this molecule, since 9·1 A is sufficiently near the 'ideal' period of 8·8 A, and the bond sequence is compatible with the chemical structure of the molecule.*

STRUCTURES CONTAINING POLYVALENT ATOMS OTHER THAN CARBON

The amount of evidence on molecules containing nitrogen, oxygen and sulphur atoms bonded with carbon is less than that relating to carbon-carbon links, but what evidence there is suggests that in these molecules also the principle of staggered bonds is obeyed. This would perhaps be expected, because these atoms are stereochemically similar to carbon, having their bonds fixed at the tetrahedral angle to each other, and therefore the interactions of the bonds of successions of any of these atoms might be expected to be similar to those of carbon sequences.

The only evidence on nitrogen compounds is that the identity periods of poly-hexamethylene adipamide $(\text{—NH(CH}_2)_6\text{.NH.CO.(CH}_2)_4\text{.CO—})_n$ (Fuller 1940), and natural silk $(\text{—CH}R\text{.CO.NH—})_n$ (Kratky & Kuriyama 1931) are such that the chains must be plane or nearly plane zigzags similar to the carbon chain in normal paraffin hydrocarbons; and the bond structure of hexamethylene tetramine, $(CH_2)_6N_4$, is the same as that of a portion of the diamond structure (Dickinson & Raymond 1923; Wyckoff & Corey 1934).

---

* Since this was written, the crystal structure of rubber hydrochloride has been determined by interpretation of X-ray diffraction patterns. The predicted molecular structure has proved to be correct.

The oxygen and sulphur compounds which have been examined are penta-erythritol tetra-acetate, $C(CH_2O.CO.CH_3)_4$, in which the singly linked oxygen atoms participate in plane zigzag chains (Goodwin & Hardy 1938); trimethylene oxide $(CH_2O)_3$ and the corresponding sulphur compound $(CH_2S)_3$ which have ring structures with chair forms like cyclohexane (Moerman 1937; Moerman & Wiebenga 1937); paraldehyde $(CH_3.CHO)_3$, which probably has the same ring structure (Carpenter & Brockway 1936). In all these, the bonds are staggered. In metaldehyde $(CH_3CHO)_4$ the molecules are non-planar eight-atom rings in which there is some deviation from the staggered bond positions. The identity periods of some of the polyesters $(\text{—CO—(CH}_2)_x\text{—CO—O—(CH}_2)_y\text{—O—})_n$ indicate that the chains are nearly fully extended zigzags like $(CH_2)_n$ (Fuller 1940); there cannot be much deviation from the staggered positions here.

Assuming that the staggered positions are favoured in such molecules, configurations for molecules containing these atoms can be predicted. The chain forms illustrated in figure 4 are valid for chains containing these atoms, though the identity periods will necessarily be different. It has been remarked (Fuller & Erickson 1937) that the identity periods of certain long-chain polyesters indicate that the chains cannot be plane zigzags; for instance, polyethylene-succinate,

$$(\text{—CO—CH}_2\text{—CH}_2\text{—CO—O—CH}_2\text{—CH}_2\text{—O})_n,$$

has an identity period of 8·3 A, whereas a plane zigzag chain would have a period of 9·8 A. The non-planar zigzag chain form which has been suggested for this molecule by Fuller & Erickson is actually *ABAC* illustrated in figure 4. (The period for the molecule is necessarily twice *ABAC* on account of the chemical constitution of the molecule.) These authors consider that this chain configuration results simply from attractions of atoms or dipoles placed at intervals along the chain. The present point of view is that the possible chain forms are determined by the principle of staggered bonds, though the choice by any particular molecule of one chain form rather than another no doubt depends on attractions or repulsions of certain atoms or groups not only in the same molecule but also in neighbouring molecules in the crystal. *ABAC* is not the only chain form possible for this molecule; $(AB)_2(AC)_2$ (figure 4) seems also eligible.

The complex chain forms of poly-oxy-methylene, $(\text{—CH}_2\text{—O—})_n$, and poly-oxy-ethylene, $(\text{—CH}_2\text{—CH}_2\text{—O—})_n$, probably also result from the operation of the principle of staggered bonds, combined with the effects of the interactions of particular atoms or dipoles. The identity periods—17·35 and 19·5 A respectively, each containing nine chemical units (Sauter 1932, 1933)—seem remarkably long for molecules of such essentially simple chemical structure; but it must be remembered that comparatively simple bond sequences may lead to long identity periods. Repetition of *ABABAC* leads to an 18-atom period $(ABABAC)_3$ of about 17·5 A; this seems a possible chain form for poly-oxy-methylene.

BOND POSITIONS IN UNSATURATED MOLECULES

This enquiry started as an attempt to understand the non-planar chain forms of three unsaturated polymer molecules; it was undertaken as a necessary step in the approach to the question of molecular flexibility. Now return to these unsaturated chains, and consider them against the background of the principle of staggered bonds. The interest centres on the positions of the single bonds $(\text{CH}_2\text{—CH}_2)$ which join isoprene or chloroprene units $(\text{—CH}_2\text{—CR═CH—CH}_2\text{—})$. But first of all the configuration of the isoprene and chloroprene units themselves needs some comment.

Each such unit is roughly planar (see figures 2, 4 and 7 of Part I). This is in accordance with expectation, since in ethylene (Thompson 1939) and its chloro substitution products (Brockway, Beach & Pauling 1935), all the atoms lie in a plane—a configuration consistent with the simple geometrical conception of the double bond shown in figure 5. There are, however, some large distortions, the greatest being those of the bonds holding methyl groups or chlorine atoms. Thus, in β gutta-percha the bond holding the methyl group lies over 20° out of the plane $\text{CH}_2\text{—C═CH}$ (see figure 6). In rubber, the distortions are different in the two isoprene units which make up the molecular identity period; the greatest is as much as 35°. In polychloroprene, whose molecular structure is analogous to that of β gutta-percha, the corresponding distortion of the C—Cl bond is even higher—40°. These distortions appear to be due chiefly to repulsion between the methyl or chlorine side-substituent and one of the chain CH_2 groups (5 and 4*b* in figure 6). But in rubber, since the configurations of the two isoprene units

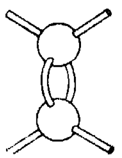
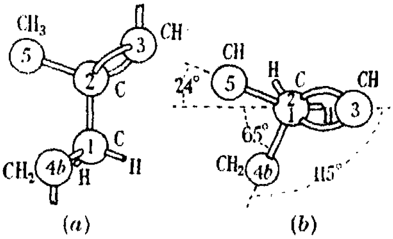


FIGURE 5. Bond positions in group $>\text{C═C}<$.

in its period are different, the distortions must be due partly to the effect of packing together the awkwardly shaped molecules; each molecule is distorted by the other molecules surrounding it; the Van der Waals' forces of attraction responsible for holding the molecules together in a precise crystalline formation are strong enough to give rise to very appreciable distortions of bond positions.

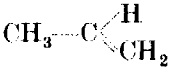
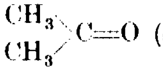
These distortions are surprisingly large. It is well established (by the great stability of *cis* and *trans* isomers) that rotation round a double bond requires a high energy—a state of affairs well illustrated by the traditional simple geometrical picture of the double bond shown in figure 5—and therefore it might have been expected that the single bond positions would be severely restricted. This is not so; evidently distortions up to 40° are fairly easily brought about.



(*a*) (*b*)

FIGURE 6. Bond positions in β gutta-percha.

Now return to the main question—the reasons for the positions of the single CH_2—CH_2 bonds which join isoprene or chloroprene units together. These bonds make large angles with the CH_2=C—CH planes, and give the chains their pronouncedly non-planar character. In β gutta-percha, plane CH_2—CH_2—C makes an angle of 115° with plane CH_2—C=CH and 89° with plane CH_2—C—CH_3 (see figure 6*b*). In polychloroprene the angles are not very different. In rubber, the two isoprene units in the period have different configurations; in one of them, the angles specified are 109 and 86°; in the other they are 169½ and 45½°. These bond positions are, no doubt, partly due to the presence of the substituent methyl group or chlorine atom; owing to its presence, the CH_2 chain group cannot be in the plane of the isoprene unit, but is pushed to one side. Thus the same repulsion which is responsible for pushing the methyl group away from the CH_2—C=CH plane (the repulsion between CH_3 group 5 and CH_2 group 4*b* in figure 6) is also responsible, at any rate partly, for the pronouncedly non-planar character of the chain. But in view of the preference for the staggered bond configuration in singly linked molecules—a preference which, it is suggested, is due largely to an orienting tendency of the bonds themselves rather than to interactions between the atoms held by the bonds—the question arises whether there is a similar influence in the unsaturated molecules under discussion. From this point of view, it is interesting to note that in β gutta-percha, if the double bond is represented in the simple geometrical way shown in figure 6, then the three bonds of atom 2 are staggered with respect to those of atom 1; this is shown most clearly in figure 6*b*. There is some distortion, but, nevertheless, an approximation to the staggered configuration, if it is justifiable to regard the double bond as two distorted single bonds. In polychloroprene, the bond positions are very similar to those in β gutta-percha. In rubber, the bond orientation for half the isoprene units is again similar, but in the other half the angles (given above) are considerably different from those in β gutta-percha. It is evident that, if we are right in supposing that there is some tendency towards the staggered bond configuration (which, together with the repulsions between atomic groups, is responsible for the non-planar character of the chain) nevertheless this configuration is easily distorted; in rubber crystals, the distortion is brought about by the Van der Waals' forces between adjacent molecules.

Ease of distortion is bound up with the magnitude of the potential barrier to rotation; there is some evidence of this magnitude in simple gas molecules, and it is interesting to find that in those containing the bond sequence —C—C=C, the barrier to rotation is much lower than in saturated molecules. Thus, in propylene,  $CH_3-C{<}^{H}_{CH_2}$ (Kistiakowsky, Lacher & Ransom 1938), and in acetone,  ${}^{CH_3}_{CH_3}{>}C{=}O$ (Schumann & Aston 1938), the barrier of rotation of the methyl group is not more than 1000 cal.—only one-third the figure for ethane and other saturated molecules. It has been suggested (Kistiakowsky *et al.* 1938) that this indicates that the double bond has no orienting influence; the rotation of each methyl group in propylene and acetone is perhaps restricted by only one single bond—the H—C bond in propylene (as compared with three in saturated molecules), and this may be the reason why the restricting potential has only one-third the value found in saturated molecules. However this may be, the smallness of the potential barrier to rotation in such groupings is presumably the reason why the staggered bond configuration is so easily distorted in rubber molecules.

In view of the experimental evidence on rubber and the thermochemical evidence just mentioned, it is to be expected that in general, molecules containing double bonds are likely to exhibit much greater deviations from the staggered bond positions than are saturated molecules.* This will not always be so, as is evident from the examples of β gutta-percha and polychloroprene; presumably in these crystals the Van der Waals' forces arising from the manner of packing of the molecules happen to favour the staggered position, whereas in rubber the packing is such that in some places the Van der Waals' forces oppose the staggered configuration.

The nomenclature already described for saturated chains may also be used for chains containing double bonds, provided that additional information on the *cis* or *trans* configuration of the double-bonded unit is included. Thus β gutta-percha (figure 2 of Part I) is (*trans*) *BAC* or (*trans*) *CAB*, α gutta-percha (figure 5*a* of Part I) is (*trans*) *BAC* (*trans*) *CAB*, and the γ form of *trans* chain (figure 5*b* of Part I) is (*trans*) *BAB* (*trans*) *CAC*. Rubber is (*cis*) *BAC* (*cis*) *BAC*. The other form of *cis* chain shown in this diagram is (*cis*) *BAC* (*cis*) *CAB*.

The discussion of ease of distortion and the magnitude of the potential barrier to rotation brings us to the question of molecular flexibility, and its connexion with rubber-like properties. This topic will be taken up in Part III of this work, but meanwhile it may be remarked that molecules containing double bonds in the chain are likely to be more flexible (other things, such as side-chain hindrances to rotation, being equal) than saturated molecules, simply because the potential barrier to rotation round the single bonds due to the orienting tendency of the bonds themselves is lower when double bonds are present in the chain than in saturated chains; and therefore, if we accept in a general way the idea that rubber-like properties are likely to be shown only by substances having very flexible molecules, then this constitutes a special reason why polymers having double bonds in their chains are likely to have rubber-like properties. Perhaps this is why so many 1 : 4 polymers of butadiene derivatives, both natural and synthetic, have rubber-like properties—simply because every fourth chain bond is a double bond. To go any further it is necessary to consider both the preferred bond positions (and the potential barriers associated with them) and the effect of side groups in increasing or decreasing the heights of the barriers to rotation. This will be done in Part III.

---

* This refers only to molecules which, like rubber, contain straightforward double bonds. Resonating systems exhibiting bonds of intermediate character are excluded from consideration.

REFERENCES

Bartholomé & Karweil 1938 *Z. phys. Chem.* B, **39**, 1.
Beach & Palmer 1938 *J. Chem. Phys.* **6**, 639.
Beach & Stevenson 1938 *J. Amer. Chem. Soc.* **60**, 475.
Beach & Turkevitch 1939 *J. Amer. Chem. Soc.* **61**, 303.
Beevers & Hughes 1941 *Proc. Roy. Soc.* A, **177**, 251.
Bragg, Sir W. H. 1913 *Proc. Roy. Soc.* A, **89**, 277.
Brockway, Beach & Pauling 1935 *J. Amer. Chem. Soc.* **57**, 2693.
Bunn 1939 *Trans. Faraday Soc.* **35**, 482.
Bunn 1941 *Proc. Roy. Soc.* (in the Press).
Carpenter & Brockway 1936 *J. Amer. Chem. Soc.* **58**, 1270.
Dickinson & Bilicke 1928 *J. Amer. Chem. Soc.* **50**, 764.
Dickinson & Raymond 1923 *J. Amer. Chem. Soc.* **45**, 22.
Ehrhardt 1932 *Phys. Z.* **33**, 605.
Eucken & Schäffer 1939 *Naturwissenschaften*, **8**, 122.
Fuller 1940 *Chem. Rev.* **26**, 143.
Fuller & Erickson 1937 *J. Amer. Chem. Soc.* **59**, 344.
Fuller, Frosch & Pape 1940 *J. Amer. Chem. Soc.* **62**, 1905.
Gehman, Field & Dinsmore 1938 *Proc. Rubber Technology Conference*, p. 961.
Goodwin & Hardy 1938 *Proc. Roy. Soc.* A, **164**, 369.
Gorin, Walter & Eyring 1939 *J. Amer. Chem. Soc.* **61**, 1876.
Kistiakowsky, Lacher & Ransom 1938 *J. Chem. Phys.* **6**, 900.
Klug 1935 *Z. Kristallogr.* **90**, 495.
Kohlrausch 1932 *Z. phys. Chem.* B, **18**, 61.
Kratky & Kuriyama 1931 *J. Phys. Chem.* B, **11**, 363.
Llewellyn, Cox & Goodwin 1937 *J. Chem. Soc.* p. 883.
Moerman 1937 *Rec. trav. chim. Pays-Bas*, **56**, 661.
Moerman & Wiebenga 1937 *Z. Kristallogr.* **97**, 323.
Pauling & Brockway 1937 *J. Amer. Chem. Soc.* **59**, 1223.
Pitzer 1937 *J. Chem. Phys.* **5**, 473.
Sauter 1932 *Z. phys. Chem.* B, **18**, 417.
Sauter 1933 *Z. phys. Chem.* B, **21**, 161, 186.
Schumann & Aston 1938 *J. Chem. Phys.* **6**, 485.
Stevenson & Schomaker 1939 *J. Amer. Chem. Soc.* **61**, 3173.
Thompson 1939 *Trans. Faraday Soc.* **35**, 697.
Turkevitch & Beach 1939 *J. Amer. Chem. Soc.* **61**, 3127.
Verweel & MacGillavry 1938 *Nature, Lond.*, **142**, 161.
Wyckoff & Corey 1934 *Z. Kristallogr.* **89**, 462.

Molecular structure and rubber-like elasticity

III. Molecular movements in rubber-like polymers

By C. W. Bunn, B.A., B.Sc.

The Research Department of Imperial Chemical Industries Ltd., Northwich

(*Communicated by Sir William Bragg, F.R.S.—Received* 12 *September* 1941)

The evidence of X-ray diffraction photographs shows that in gutta-percha and in polychloroprene at the 'melting'-point, major changes of molecular configuration due to rotation round the single bonds occur spontaneously. Melting is ascribed to the onset of chain-bond oscillation between alternative, geometrically equivalent positions ('molecular wriggling'), and the low melting-points of these substances (and also, by analogy, of rubber) as compared with polyethylene are ascribed to the greater freedom of rotation round the single bonds in the former substances.

It is postulated that the energy required for rotation round the single bonds is made up of two factors—the bond-orientation energy, and the interaction of the atoms or groups held by the bonds. The comparative freedom of rotation in rubber and polychloroprene is due to two causes: first, the presence of double bonds in the chains, which has the effect of making rotation round adjacent single bonds more free than in comparable saturated molecules; and secondly, the small hindrance to rotation offered by the side substituents. In gutta-percha the bond-orientation energy is presumably the same as in rubber and polychloroprene, but geometrical hindrance is greater, and it is suggested that this is the reason for its rather higher melting-point.

The structure of amorphous rubber is considered. Starting from the known structures of crystalline polymers, which consist of crystals tied together by molecules each of which passes through a number of crystals, the minimum change on melting is assumed, and that change—the wriggling and consequent disarrangement of the molecules—is well founded on experimental evidence; it is not necessary to assume that the molecules have a completely random configuration, and there are reasons why this is unlikely. The mechanical properties above and below the melting-point are correlated on this basis. The statistical effects to which Mark has drawn attention operate through the 'tying' portions of the molecules.

In Part I of this work (1941), the crystal structures of β gutta-percha, rubber and polychloroprene (deduced from X-ray diffraction photographs) were described. In pursuance of the idea that rubber-like properties are due to molecular flexibility resulting from the swivelling of the chain units round single chain bonds, it is necessary to consider which chain-bond positions are the most stable and what hindrances there are to rotation away from these positions. The question of the most stable bond positions was considered in Part II (1941). I now consider the evidence for the occurrence of rotation in rubber-like substances, the hindrances to rotation in different molecules and their effect on the crystallization or 'melting' temperature, and the explanation of the mechanical properties of these substances in terms of structure and molecular movement, both above and below the crystallization temperature.

Rubber is non-crystalline and elastic at room temperature; but on cooling below 0° C it crystallizes (Clark, Wolthuis & Smith 1937; Clark, Gross & Smith 1939), and in that condition has the mechanical properties of gutta-percha (Houwink 1937); that is to say, a frozen specimen no longer has enormous elasticity, but it can be drawn or rolled out irreversibly, the crystals becoming oriented. Conversely, if gutta-percha is warmed to 70° C, it becomes amorphous (transparent, optically isotropic and non-crystalline), and its mechanical properties are then like those of rubber—it is soft and elastic (Hauser & Susich 1931). The difference between the two substances thus appears to be essentially a difference of crystallization temperature. These crystallization temperatures are remarkably low; polyethylene, $(—CH_2—CH_2—)_n$, of comparable molecular length crystallizes at 115–125° C. In attempting to understand rubber-like elasticity in terms of molecular behaviour, the first question is why are the crystallization points of such enormous molecules so low, and why is the crystallization point of the *cis* form of polyisoprene (rubber) lower than that of the *trans* form (gutta-percha).

THE CRYSTALLIZATION POINTS OF LONG-CHAIN POLYMERS

When a crystalline solid is heated, it melts when the amplitude of the thermal motions of the atoms, ions or molecules of which it is composed becomes great enough to destroy the regular arrangement in the crystals. The Lindemann equation (Lindemann 1910) embodies the assumption that melting occurs when the amplitude of vibration of the structural units becomes equal to the distances between these units, and for a few metallic elements and simple salts it has been shown (see Eucken 1930) that this equation leads to vibration frequencies in agreement with those calculated in other ways (from specific heats, 'reststrahlen', and compressibilities). Similar calculations for molecular crystals have not been made, but it is certain that their melting-points are determined by some such condition; the general rise in melting-point with the molecular weight of organic molecules testifies to this, since the amplitude of vibration of a molecule at a given temperature must decrease with rising molecular weight. But this is valid only for rigid molecules; when one part of a molecule can make large movements with respect to another part, the mass of the moving parts, which controls the melting-point, is not the mass of the whole molecule but that of the comparatively rigid portions of the molecule which move with respect to each other. Crystals composed of very flexible molecules are expected to have melting-points much lower than those composed of rigid molecules of the same weight.

The molecules of long-chain polymers with which we are concerned consist of fairly rigid units joined by single bonds round which rotation can occur relatively easily. For instance, in a rubber molecule each isoprene unit $-CH_2-C(CH)_3=CH-CH_2-$ is relatively rigid owing to the presence of the double bond, but rotation can occur round the single bonds which join the isoprene units together. In crystals of chain polymers, the molecules are packed together like a bundle of rods; vibrations perpendicular to the rods are likely to be very much larger than those along the rods, and it is therefore the former which cause melting when their amplitude reaches the required magnitude. The fact that the melting-point of polyethylene is practically independent of the molecular weight when the latter is above 10,000* is in agreement with this conception. These lateral movements may be due partly to distortion of the bond angles, and partly to rotation round the single bonds; in some polymer molecules, rotation round the single bonds through large angles may occur—for example, through about 120° to alternative bond positions, or even through 360°; the amplitude of movement occurring in this way is far greater than that likely to result from bond-angle distortion, and therefore, ease of rotation round the single bonds is likely to be the main influence determining the melting-point, at any rate in hydrocarbon polymers, in which the intermolecular forces are comparatively weak. The melting-points of crystals composed of molecules in which rotation is easy are likely to be many degrees lower than those composed of molecules in which rotation does not occur.

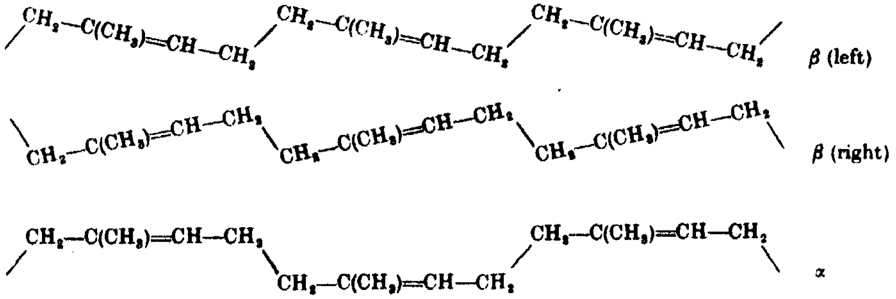
First of all, evidence will be presented which leaves no doubt that major changes of molecular configuration due to rotation round the single bonds do occur in gutta-percha and polychloroprene at the crystallization point.

EVIDENCE ON MOLECULAR MOVEMENTS AT THE CRYSTALLIZATION POINT

Gutta-percha at 70° C is amorphous. On cooling, it may crystallize in two different forms, known as α (formed on slow cooling) and β (formed on rapid cooling). These two forms are not simply different arrangements of the same molecules, but have different chain structures; the identity periods of the chains (4·7 A for the β form produced by rapid cooling, 8·9 A for the α form produced by slow cooling) leave no doubt of this. The difference between the two chains is simply a difference in the position of the single bonds (CH_2—CH_2) which join the (*trans*) isoprene units together (see Part I). The two types of molecules are illustrated in figures 2 and 5*a* of Part I; they may be represented thus:

---

* Unpublished work of this laboratory.



(Each isoprene unit is, as far as its carbon atoms are concerned, approximately planar, the plane being perpendicular to the paper; the connecting links—CH_2—CH_2 bonds—lie nearly *in* the plane of the paper.)

Evidently, the single bonds which determine the molecular configuration switch over spontaneously from one position to another in the amorphous material.

That the change of molecular configuration occurs at the moment when a molecule takes its place in a crystal is confirmed by the fact that in any one β crystal the molecules are all either left-handed or right-handed, not mixed. For, suppose β molecules were formed previous to crystallization; the probability of formation is the same for left-handed as for right-handed molecules; hence in any group of molecules we should expect to find equal numbers of the two types, and when the group crystallized we should get an arrangement containing equal numbers of left- and right-handed molecules. (It is very unlikely that, in a non-fluid medium, the molecules would be able to sort themselves out into separate groups of thousands of left-handed molecules and thousands of right-handed molecules. Limited movements, necessary to convert a random grouping into an ordered arrangement, must occur, but extensive migrations of large numbers of molecules are unlikely.) It is thus necessary to assume that, at the moment of crystallization, each molecule assumes whichever configuration is necessary for convenient packing with other surrounding molecules. Evidently, the most stable mode of packing is one in which all the molecules in one crystal have the same configuration, and accordingly, molecules change their configuration as required. Conversely, when a crystal 'melts', bond oscillation sets in at the melting-point.

The same argument with regard to change of configuration of β-type molecules applies to polychloroprene, which is structurally analogous to β gutta-percha. For rubber itself we have no evidence on the type of molecular movement which occurs, since its crystal structure happens to include both left- and right-handed molecules; it is, however, reasonable to suppose that in rubber the type of molecular movement is similar to that in polychloroprene and gutta-percha at or above their melting-points, since the molecules of all these substances are built on the same principles (see Part I).

'Free rotation' round simple bonds is, of course, a well-known phenomenon in liquids and solutions, and the conception has been much used in recent theories of the structure of rubber-like polymers; but the evidence now presented is the first direct evidence that such rotation does occur spontaneously in rubber-like polymers. The term 'free rotation' is perhaps a little misleading, as it suggests rotation at uniform speed, or quite random rotatory movements. Actually, in all probability, there is a rapid switching of the bonds from one definite position to another, as in ethane and its derivatives. (See Part II of this work.) This is probably what happens in the rubber-like polymers discussed here; the CH_2—CH_2 chain bonds are continually oscillating from one definite position to the other. When this happens in a long-chain polymer molecule, it means that the whole molecule is spontaneously wriggling; an attempt to suggest this is made in figure 1, in which the molecules are represented in skeletal fashion to avoid confusion. Two modes of wriggling, both consistent with the maintenance of essentially straight chains, are illustrated, but many other modes are possible. The movements may not be confined to oscillation between two positions; the third bond position allowed by the principle of staggered bonds (see Part II) may also be visited; there is, however, no evidence on this point.

I now ask whether the onset of molecular wriggling is directly responsible for the melting of the crystals, or whether such wriggling can occur in the crystals. It is necessary to ask this question, because it appears that in certain crystals whole molecules rotate; in other words, the rotation of the molecules does not cause melting. Now molecular rotation at temperatures far below the melting-point only occurs in crystals composed of molecules which are roughly spherical (e.g. camphor derivatives) or disk-shaped (White & Bishop 1940; White, Biggs & Morgan 1940). These molecules can rotate without disturbing their neighbours unduly, and this is evidently the reason why the rotation of the molecules does not cause the crystals to melt. Among chain compounds it appears that in some crystals of normal paraffins and alcohols the molecules rotate about their long axes (causing a phase change) a few degrees below the melting-point (Bernal 1932; Müller 1932). Such molecules have roughly cylindrical symmetry, yet it appears that when a molecule, through its vibrations, pushes away its neighbours to an extent sufficient to permit complete rotation of the molecule, the amplitude of movement is nearly sufficient to melt the crystal; a further rise of temperature of only a few degrees is necessary to cause melting. Now the amplitude of movement in gutta-percha molecules which are wriggling in the manner described is much greater than that necessary for rotation of the roughly cylindrical molecules just mentioned. Hence, it seems reasonable to suppose that the wriggling is not likely to occur in the crystals, but that the onset of molecular wriggling is directly responsible for the melting of the crystals. There is no evidence on the matter; the X-ray patterns which indicate a precise structure in the crystals (see Part I) do not rule out the possibility of wriggling in the crystals, for changes of configuration might take only a small fraction of the time, and thus might not affect the X-ray patterns; but it certainly seems unlikely. At any rate, it will be assumed as a working hypothesis that the onset of molecular wriggling is responsible for the melting of the crystals.

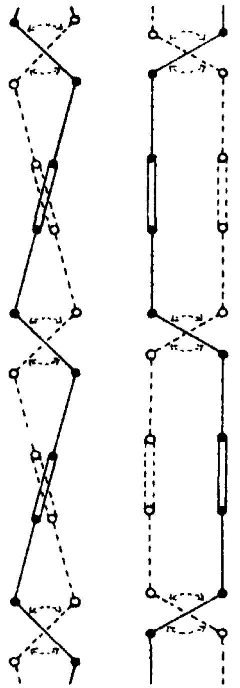
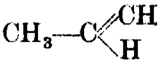
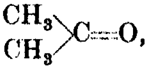


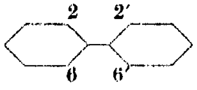
FIGURE 1. Gutta-percha molecules wriggling.

On this hypothesis, the principal reason why the crystallization temperatures of rubber and similar substances are so remarkably low is that rotation round the single bonds is relatively easy (compared with polyethylene, for instance). Intramolecular forces are so free that the Van der Waals forces urging crystallization are balanced at a much lower temperature than that to be expected for more rigid molecules of similar size. In other words, each isoprene (or chloroprene) unit has a freedom of movement approaching that of monomeric isoprene or chloroprene; hence the crystallization points of polyisoprene and polychloroprene approach those of the small monomer molecules. Other effects which might influence the crystallization point are the packing difficulties of awkwardly shaped molecules; if molecules do not pack neatly the Van der Waals forces cannot play their full part. The crystal densities give some indication of the efficiency of packing; those of gutta-percha, rubber and polyethylene, calculated from the unit cell dimensions of the crystals, are 1·04, 1·00 and 1·00 respectively; these are so similar that the influence of this factor on the crystallization point is likely to be quite subsidiary.

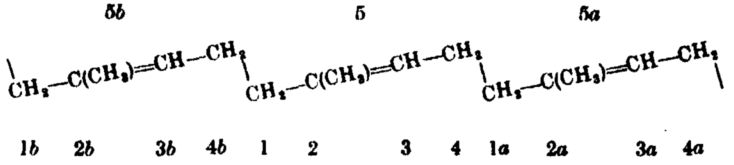
If accepted, the hypothesis that chain-bond oscillation (causing major changes of molecular configuration) is responsible for the melting of these rubber-like polymers leads to the inquiry why rotation round the single bonds occurs at a lower temperature in rubber and polychloroprene than in gutta-percha; and further, why the crystallization temperatures of all three substances are so much lower than that of polyethylene.

Hindered rotation in chain polymers

It will be assumed that the most stable positions for the bonds, and the energy required to move a bond from one position to another, are determined by two factors: the orienting tendency of the electron pairs forming the bonds, and the interactions of the atoms or groups held by the bonds. It has been suggested that the potential barrier to rotation in ethane and propane—some 3000 cal. per methyl group—is chiefly due to the first of these factors, the orienting tendency of the bonds themselves (Kistiakowsky, Lacher & Ransom 1938). It will be assumed, then, that in a saturated molecule, bond orientation alone is responsible for a barrier of this order of magnitude. Now in propylene,  $CH_3{-}C\begin{smallmatrix}\diagup CH_2\\ \diagdown H\end{smallmatrix}$, and in acetone,  $\begin{smallmatrix}CH_3\diagdown\\ CH_3\diagup\end{smallmatrix}C{=}O$, the methyl groups rotate more easily; the barrier is 1000–2000 cal. (Kistiakowsky *et al.* 1938; Crawford *et al.* 1939; Schumann & Aston 1938), apparently owing to the presence of the double bond. Presumably the same effects occur in chain polymer molecules; when a double bond is present, as in rubber, gutta-percha and polychloroprene, rotation round the adjacent single bond is likely to occur more easily than in a comparable saturated molecule. This may be the reason why the melting-points of these polymers which have double bonds in their chains are low.

The explanation of the differences between the three unsaturated polymers is to be sought in the other factor—the interactions of the groups held by the bonds. Evidence relevant to this question is available in the large volume of work which has been done on the racemization of diphenyl derivatives (Adams & Yuan 1933). It appears that for these substances the predominant factor controlling ease of rotation round the single bond is the size of substituent groups in key positions. Thus, when fluorine atoms are present at the 2, 2′, 6 and 6′ positions of diphenyl,  racemization (due to rotation round the single bond) is so easy that resolution into optically active forms cannot be achieved, while if the larger chlorine atoms or methyl groups are present in these positions, resolution can be effected and racemization does not occur in solution even at 100° C. The polarity of groups in these positions, and the presence of substituents in other positions, have only secondary effects. Therefore, in considering ease of rotation in the three unsaturated polymers, I look for geometrical hindrances to rotation.

In gutta-percha, it is known that at the melting-point



$$\backslash CH_2{-}C(CH_3){=}CH{-}CH_2\backslash CH_2{-}C(CH_3){=}CH{-}CH_2\backslash CH_2{-}C(CH_3){=}CH{-}CH_2\backslash$$

the CH_2—CH_2 chain bonds (4*b*–1 and 4–1*a*) swing from one side of the isoprene unit to the other. The shortest route from one position to the other is illustrated in figure 2; bond 1–4*b* swings (round bond 1–2 as axis) through about 120°. When this happens, carbon atom 4*b* has to pass within 2·6A of the methyl carbon atom 5; since the external radius of a CH_2 or a CH_3 group is 1·8–2·05A, there will evidently be some hindrance to this rotation. (The sizes of the balls representing atoms in figure 2 are arbitrary; actually the electron clouds are much larger, so much so that joined atoms overlap enormously. But to give the electron clouds their correct sizes would obscure the stereochemistry; therefore in the drawing the atoms are symbolized by small balls.) There is an alternative route involving rotation through 240° in the opposite direction, but this offers no advantages; atom 1*a* would have to pass within 2·5A of 5 and 1 within 2·5A of 5*b*; and in addition, the wider sweep would be hindered by surrounding molecules. The shorter route involving rotation through 120° is therefore the more probable. It has been seen that this rotation is hindered by the size of the groups which have to pass each other, and at low temperatures no rotation is possible; but with rise of temperature the vibrations of the atoms (involving bending and stretching of the bonds) increase, until a point is reached at which rotation can occur at certain phases of the vibrations. The thermal energy must be enough to overcome (1) the small potential barrier due to preferred bond orientation, (2) the hindrance due to the size of the groups which have to pass each other, and (3) the hindrance due to other surrounding molecules. The temperature at which this occurs is about 60° C (actually 65° for the α form and 56° for the β form).

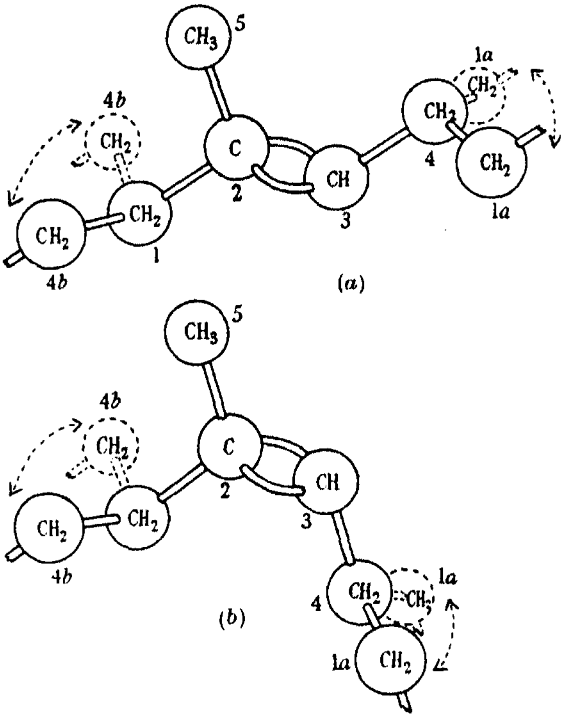


FIGURE 2. Bond oscillation in gutta-percha (above) and rubber (below).

The molecular structure of polychloroprene is entirely analogous to that of β gutta-percha; it has a chlorine atom in place of the methyl group of gutta-percha. Now the distance between linked chlorine and carbon atoms (1·77 A) is greater than the distance between linked carbon atoms (1·53 A), and thus the chlorine atom of polychloroprene stands farther from the chain than the methyl group of gutta-percha; moreover, the external radius of a chlorine atom appears to be slightly less than that of a methyl group. On both accounts the chlorine atom would be expected to offer less obstruction to chain-bond rotation than the methyl group (see figure 2a). Chain-bond rotation, and thus molecular wriggling, could therefore occur at a lower temperature in polychloroprene than in gutta-percha. This may be the reason why the melting-point of polychloroprene is lower than that of gutta-percha. There are other influences—the larger inertia of the chlorine atom, and the interaction of C—Cl dipoles—which might be expected to have the opposite effect and tend to increase the melting-point; but evidently such effects are secondary, possibly because melting can be achieved without moving the chlorine atoms at all; one can imagine the chloroprene units oscillating about the C—Cl bonds, so that the chains are wriggling (and the crystals melting) while the chlorine atoms remain stationary.

In rubber, as in gutta-percha, CH_2 group *4b* (figure *2b*), in swinging round to the other side of the isoprene unit, has to get past the obstruction offered by the methyl side group. This methyl side group, however, is set slightly farther back (away from group *4b*) in rubber than in gutta-percha. (In half the isoprene units in a rubber crystal the angle CH_3—C═CH is 115°, while in gutta-percha the corresponding angle is 125°.) Furthermore, it can vibrate farther in the direction away from *4b*. In gutta-percha, group 4, only 3·0 A away at the other end of the double bond (figure *2a*), obstructs vibration in this direction, but in rubber (figure *2b*), owing to the *cis* double-bond configuration, no such obstruction is present, only a hydrogen atom being attached to C_3 on this side. Since vibration of the methyl group in the direction away from *4b* is easier, group *4b* can swing round past the methyl group at lower temperatures in rubber than in gutta-percha. This is evidently the reason why the crystallization point of rubber is lower than that of gutta-percha.

A qualitative explanation of the fact that rubber and polychloroprene are non-crystalline and elastic at lower temperatures than gutta-percha can thus be given on the basis of the simple stereochemical conception of geometrical hindrance to bond oscillation.

Comparison of these substances with saturated chain polymers is less simple, because there are two factors to be taken into account—the bond-orientation effect as well as geometrical hindrance. To form some idea of the effect on the melting-point of the higher bond-orientation barrier in saturated as compared with unsaturated molecules, we need information on the melting-points of saturated and unsaturated molecules in which the geometrical hindrances are the same. Rubber hydrochloride $(—CH_2—CH_2—CH_2—C(CH_3)Cl—)_n$ is being studied with this end in view. Meanwhile, in attempting to compare rubber and gutta-percha with polyethylene $(—CH_2—CH_2—)_n$, in addition to the difference in bond-orientation energies there is another interesting difference. In rubber and gutta-percha, bond oscillation which is apparently sufficient to cause melting occurs between geometrically enantiomorphous, energetically equivalent positions (see figure 2). But in polyethylene, the most stable form of chain is a plane zigzag, in which each chain bond is in a geometrically unique position; movement of a chain bond to either of the other two staggered positions is movement to a geometrically non-equivalent position of higher energy (higher owing to repulsion between the CH_2 groups (see Part II)). This must increase the height of the energy barrier to be surmounted. (The magnitude of this effect is unknown; potential energy diagrams for *n*-butane, similar to those given for halogen derivatives of ethane (Beach & Palmer 1938; Beach & Turkevitch 1939), would give some idea of it, but such diagrams are not available.) Any increase in the height of the energy barrier would increase molecular rigidity and therefore raise the melting-point.

In polyethylene, the bond-orientation effect and the repulsion between the CH_2 groups co-operate in favour of the stability of the plane zigzag type of carbon chain, and the co-operation of the two factors is no doubt responsible for the high melting-point. In some saturated polymer molecules having large substituent groups close together along the chain, it may happen that the two factors are in opposition: if mutual repulsion of the substituent groups opposes the staggered configuration preferred by the bonds themselves, it may well happen that the energy barrier to rotation is reduced and may even be negligible. This is probably why poly*iso*butylene $(-CH_2-C(CH_3)_2-)_n$, with its pairs of methyl groups on alternate carbon atoms (Fuller, Frosch & Pape 1940), has a low melting-point, being amorphous and rubber-like at room temperature: examination of models of this molecule shows that, no matter what chain structure is assumed, mutual repulsion between the methyl groups opposes the staggered configuration; that is to say, the bonds prefer one configuration (with the bonds staggered), while the methyl side groups prefer another, the net result being that all configurations have much the same energy, the barriers to rotation are small and the melting-point low.

Thus it seems possible to understand, in a rough general way, the melting-points of long-chain polymers on the basis of the two hypotheses assumed in the foregoing discussion. These hypotheses are: (1) that melting is due to lateral movements of segments of more or less flexible molecules, and (2) that the energy required for rotation round the single bonds (the chief basis of molecular flexibility) is made up of two factors—the bond-orientation energy, and the interaction of the atoms or groups held by the bonds.

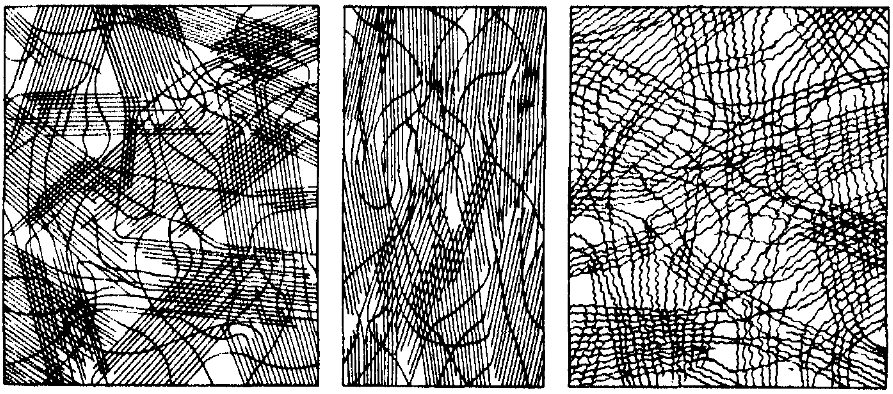
STRUCTURE OF AMORPHOUS RUBBER

I shall accept the conclusion that in amorphous gutta-percha, polychloroprene, and by analogy in rubber and other similar substances above their crystallization points, molecular wriggling (by which is meant spontaneous major change of molecular configuration of the type previously discussed) occurs spontaneously. I now inquire to what extent this leads to curling up of the molecules. It should be noted that the low melting-points of these substances can be adequately explained without assuming any curling at all; one can quite well imagine the molecules wriggling sufficiently to cause melting and yet retaining the straight chain structure found in the crystals (see figure 1). It has, however, been pointed out by Mark and others (see, for instance, Mark 1939) that in the case of an isolated molecule the very fact that the number of different curled-up configurations is much greater than the number of straight configurations must tend to favour an average curled-up configuration; in other words, each isolated molecule, owing to thermal movements, assumes a random configuration, the probability of an approximation to a straight configuration being very small. The question is to what extent these statistical effects operate in rubber-like substances, in which the molecules, so far from being isolated, are packed close together. The X-ray pattern of amorphous rubber does not determine whether the molecules are curled up or not (Simard & Warren 1936).

A theory of rubber-like elasticity was built, chiefly by Mark (1939) and Kuhn (1936), on the two assumptions that rotation round the single bonds is quite free, and that the molecules behave as if in isolation. The second assumption seemed from the outset improbable; however, a calculation by Kuhn (1936) of the modulus of elasticity of rubber gave an approximately correct figure. This seemed encouraging support for the theory; but it has since been shown by Bresler & Frenkel (1939) that if the existence of an appreciable barrier to rotation is taken into account, such calculations give moduli much smaller than the experimental values; a barrier of 3600 cal. leads to a modulus of elasticity one-twentieth the experimental value. (It must be pointed out that the figure of 3600 cal. for the height of the barrier to rotation, though possibly of the right order of magnitude, is not necessarily correct for rubber. Bresler & Frenkel's procedure was to consider the chain $(\text{—CH}_2\text{—CH}_2\text{—})_n$ and assume that the barrier has the magnitude 3600 cal. found for propane; actually, the value for propane (where a methyl group rotates but does not move as a whole in relation to the rest of the molecule) is not appropriate for the long

CH_2 chain, for which the barrier for *n*-butane (not yet known) would be more nearly correct, since in this molecule a whole CH_3 group swings round, and this would be comparable to the CH_2 group in the long chain. But, apart from this, the choice of the simple CH_2 chain as an example—a practice followed in most theoretical discussions—is an unfortunate one, for polyethylene is not a typical rubber-like substance; it is crystalline up to 115–125° C, and above the melting-point is a very viscous liquid. The magnitude of the barrier for rubber is unknown, but it is certainly not zero, and Bresler & Frenkel's calculations show that the existence of a barrier of appreciable magnitude leads to a considerable reduction in the value of the modulus calculated by Kuhn's method.) This must mean that Kuhn's second assumption—that the molecules behave as if in isolation—is not justified; the correct result of his calculation was the fortuitous outcome of making two incorrect assumptions which happened to have roughly equal and opposite effects on the result.

I conclude, therefore, that intermolecular action in rubber-like substances is by no means negligible. Its effect on the form and packing of the molecules naturally depends on the nature of the substance (the geometry of the molecules and the forces between individual atoms); but we may form some idea of the general effect of such forces by referring to the structures of ordinary liquids composed of small molecules. It appears that in such substances, disorder is not complete; there is, especially just above the melting-point, some attempt at an orderly arrangement, the type of arrangement being usually a rough approximation to that which exists in the crystal of the substance in question. In other words, although the molecules are moving about, they are more often than not surrounded by neighbouring molecules arranged in roughly the same manner as they are in the crystal (Randall 1937). For molecules of roughly spherical shape, this tendency is greatest for polar liquids such as water and least for non-polar liquids such as hydrocarbons; the shape of the molecules plays an important part only when there is considerable departure from spherical shape (see Bernal 1937). In long-chain polymers the most extreme departure from spherical shape occurs; consequently one expects to find, in these substances just above their melting-points, partially ordered arrangements of molecules corresponding roughly with the arrangements found in the crystals. In all crystals of long-chain polymers, as far as is known, the molecules are arranged parallel to each other as in a bundle of rods. (Any other arrangement would give very poor packing and a low density.) But the molecules are much larger than the crystals, and therefore any one molecule in the course of its length passes through several crystals (see Mark 1940). I therefore visualize a structure of the type shown in figure 3*a*, in which crystals (groups of straight parallel lines representing molecules) are tied together by molecules, each of which passes through a number of crystals. It is this feature which is responsible for the cohesion and high tensile strength of long-chain polymers as compared with molecular crystals composed of small molecules. Just above the melting-point, therefore, one should find that portions of molecules are associated in roughly parallel groups, as in figure 3*c*, in which the wavy lines represent molecules wriggling in the way described previously. This arrangement differs from that in the crystal in that the molecules are not precisely parallel (one can imagine them pulled away from parallel positions by the 'tying' portions of molecules) and are not arranged side by side in a regular way. With rise of temperature, the arrangement should become more and more disordered, and the molecular configuration more irregular.



(*a*) (*b*) (*c*)

FIGURE 3. (*a*) Crystalline long-chain polymer. Lines represent molecules, groups of parallel straight lines represent crystals. (*b*) The same cold-drawn. Crystals apparently parallel. (*c*) Amorphous rubber-like polymer. Wavy lines represent wriggling molecules. When rubber is frozen, the fine structure becomes like *a*; when it is stretched, the fine structure becomes like *b*.

I now inquire whether the size of the semi-ordered regions which exist just above the melting-point is at all comparable with that of the crystals which exist below the melting-point. Judging from the breadth of the diffraction spots, the crystals in the polymers mentioned in this paper are at least several hundred A across, that is to say, there are straight portions of molecules several hundred atoms long, packed together in thousands to form precise arrangements. On this question, comparison of the extensibilities of crystalline and rubber-like polymers is instructive.

MECHANICAL PROPERTIES ABOVE AND BELOW THE MELTING-POINT

It is worth noting, first of all, that the conception of molecular wriggling gives an immediately convincing reason for the characteristic semi-liquid properties of rubber-like substances—their softness, flexibility, and ease and rapidity of extension and contraction. These substances have such properties simply because the spontaneous thermal movements of segments of molecules are of similar magnitude to the movements of whole molecules in ordinary liquids, from which rubber-like substances differ in their lack of fluidity which is obviously due to the enormous length of the molecules and the resultant entangling effects. Thermal movements of such magnitude are really implied in the statistical theory of Mark & Kuhn, and in the treatment of K. H. Meyer (1939); but such movements have not received the attention they deserve. For instance, in the statistical theory the emphasis has always been placed on the conception of irregular molecular configuration in unstretched rubber, rather than on spontaneous changes of configuration; and when Meyer (1938) speaks of rubber-like substances as being 'liquid in two dimensions and solid in the third', he is referring to the forces between different molecules and between parts of the same molecule, rather than to molecular movements. The present point of view is that neither irregularity of configuration nor any particular distribution of intermolecular forces is sufficient to confer rubber-like properties. There are many amorphous polymers (especially interpolymers) in which the molecules probably have an irregular configuration, but they will not be rubber-like unless the molecules are so flexible that the thermal motions of segments of molecules cause spontaneous major changes of configuration.

As regards extensibility, it seems at first sight that there is a fundamental difference between rubber-like and crystalline specimens. The latter (for instance, polyethylene, gutta-percha and rubber hydrochloride at room temperature) can be drawn out to 3–5 times their original length, but when drawn they remain extended with the crystals all nearly parallel to each other (figure 3*b*); rubber, polychloroprene and poly*iso*butylene can be stretched 7–10 times their original length, and when released, contract spontaneously. It is found, however, that cold-drawn crystalline polymers, when warmed to temperatures a little below their melting-points, contract spontaneously and almost completely. (For published figures on rubber hydrochloride, see Gehman, Field & Dinsmore 1938.) The fact that contraction occurs when the specimens are still largely crystalline shows that a structure composed of crystalline blocks joined by 'tie' molecules can have a reversible extensibility of at any rate 400 %. The difference between this behaviour and that of rubber-like substances is only one of degree, and this suggests that the dimensions of the molecular bundles in rubber-like substances at room temperature may be not very different from those of the crystals formed below the melting-point.

In the extension and contraction of crystalline polymers, the 'tying' portions of molecules must play the most important part; the statistical effects considered by Mark, operating on the 'tying' portions of molecules, are no doubt responsible for the spontaneous contraction just below the melting-point. The crystals presumably remain, for the most part, as organized units; but it must often happen that in cold-drawing, different parts of the same crystal are dragged in different directions, and the crystal will then be sheared along cleavage planes. That this actually happens in polyethylene is shown by the fact that when orientation in a cold-drawn specimen is particularly good, the crystal size is smaller than in the original undrawn specimen (as shown by the breadth of X-ray diffractions). Similarly, in spontaneous contraction just below the melting-point, there is X-ray evidence that the size of the crystals increases; small crystals evidently join together to form larger ones. These facts all help us to understand what happens when a rubber-like substance is stretched. First of all, the molecular bundles can be figured as being pulled by the 'tying' portions of molecules into approximately parallel orientation; but slipping of individual molecules and shearing of molecular bundles into two or more groups occur very much more easily than in a crystalline polymer owing to the spontaneous wriggling of the molecules. (If a molecule gets into a jam when in one configuration, it may be able to get past easily when in another configuration.) There is consequently much more slipping and changing from one molecular bundle to another than there is in crystalline polymers; to this may be ascribed not only the far greater ease of extension, but also the greater extensibility of rubber-like as compared with crystalline polymers.

CRYSTALLIZATION BY STRETCHING

During the stretching of rubber, crystals are formed and the structure of fully stretched rubber is analogous to that of cold-drawn crystalline polymers (figure 3*b*). Crystallization has usually been attributed to the straightening out of curled-up molecules, which, when they find themselves parallel to each other, are able to form a precise arrangement; but there is another effect which may be of equal or greater importance. If a tangle of string is pulled out as if it were a piece of rubber, not only do many of the strands become roughly parallel to the direction of pulling, but also it is obvious that many strands are laterally compressed by surrounding strands. Therefore, in stretched rubber, not only are the bundles of molecules roughly parallel, but many of them are in a state of lateral compression. The molecules are forced closer together laterally, and are obliged to take up less space; the obvious way of taking up less space laterally is to stop wriggling and pack together neatly; this the molecules do, forming crystals. The greater the extension, the greater the proportion of molecular bundles laterally compressed, and therefore the greater the proportion of crystals.

(It is necessary to mention that, although the properties of gutta-percha above its melting-point (65° C) are in general similar to those of rubber at room temperature, it does not crystallize on stretching (Hauser & Susich 1931). The reason for this is probably that gutta-percha molecules are considerably shorter than rubber molecules; there is more internal irreversible slipping, as is shown by the fact that stretched amorphous gutta-percha on release does not return to its original dimensions. Consequently it is not possible to get sufficient tension on the specimen to cause crystallization at 70° C.)

In mechanical properties, the greatest contrast between rubber-like and crystalline polymers is shown by the behaviour of stretched specimens on release. Polyethylene and gutta-percha at room temperature remain extended (except for a small contraction of 10 %) and crystalline, while rubber contracts spontaneously and almost completely and becomes amorphous. The reason why the former substances remain extended is that, during stretching, the crystals are to some extent pulled to pieces and put together again in a different way. On release, the statistical effects to which Mark has drawn attention tend to make the 'tying' portions of molecules assume a random configuration; this would disorient the crystals and thus cause contraction to the original length. But contraction cannot go very far unless the jammed-in crystals can be pulled to pieces by shearing along cleavage planes; the pull of the 'tie' molecules is not sufficient to achieve this, and the specimen therefore remains extended except for a small contraction of 10 %. (Near the melting-point, shearing becomes easier and the increased thermal motions of the 'tie' molecules are sufficient to pull the crystals to pieces, and spontaneous contraction occurs.) In contrast, when stretched (largely crystalline) rubber is released, a small initial contraction leads to a slight relaxation of lateral compression; the crystals now have more room, and molecules on the surfaces of some crystals start wriggling. This loosening of the structure then permits a further contraction; more molecules then wriggle, and the process spreads through the whole specimen, until the 'melting' of the crystals and the contraction of the specimen is complete. The rapidity of contraction ('snap') is due to the fact that wriggling molecules slip past each other almost as easily as the molecules in an ordinary liquid.

A reasonable general account of the mechanical properties of crystalline and rubber-like polymers and the relation of these types to each other can thus be given on the basis of the conceptions developed in the foregoing discussion.

I wish to thank Mr J. L. Matthews of this laboratory for his assistance in the experimental work on which Part I of this paper was based. I am also indebted to Professor C. N. Hinshelwood, F.R.S., Dr W. T. Astbury, F.R.S., Dr R. E. Slade, Dr W. J. S. Naunton, and many of my colleagues in this laboratory for advice, encouragement and criticism.

REFERENCES

Adams & Yuan 1933 *Chem. Rev.* **12**, 261.
Beach & Palmer 1938 *J. Chem. Phys.* **6**, 639.
Beach & Turkevitch 1939 *J. Amer. Chem. Soc.* **61**, 303.
Bernal 1932 *Z. Kristallogr.* **83**, 153.
Bernal 1937 *Trans. Faraday Soc.* **33**, 27.
Bresler & Frenkel 1939 *Acta Physicochim. U.R.S.S.* **11**, 485.
Bunn 1941 *Proc. Roy. Soc.* (in the Press).
Clark, Gross & Smith 1939 *Bur. Stand. J. Res.* **23**, 1.
Clark, Wolthuis & Smith 1937 *Bur. Stand. J. Res.* **19**, 479.
Crawford, Kistiakowsky, Rice, Wells & Wilson 1939 *J. Amer. Chem. Soc.* **61**, 2980.
Eucken 1930 *Lehrbuch der Chemischen Physik*, p. 156.
Fuller, Frosch & Pape 1940 *J. Amer. Chem. Soc.* **62**, 1905.
Gehman, Field & Dinsmore 1938 *Proc. Rubber Techn. Conf.* p. 961.
Hauser & Susich 1931 *Kautschuk*, **7**, 120, 125, 145.
Houwink 1937 *Elasticity, Plasticity and Structure of Matter*, p. 231.
Kistiakowsky, Lacher & Ransom 1938 *J. Chem. Phys.* **6**, 900.
Kuhn 1936 *Kolloidzschr.* **76**, 258.
Lindemann 1910 *Phys. Z.* **11**, 609.
Mark 1939 *Chem. Rev.* **25**, 121.
Mark 1940 *J. Phys. Chem.* **44**, 764.
Meyer 1938 *Chem. Ind. Rev.* **57**, 439.
Meyer 1939 *Chem. Rev.* **25**, 137.
Müller 1932 *Proc. Roy. Soc.* **138**, 514.
Randall 1937 *Trans. Faraday Soc.* **33**, 105.
Schumann & Aston 1938 *J. Chem. Phys.* **6**, 485.
Simard & Warren 1936 *J. Amer. Chem. Soc.* **58**, 507.
White, Biggs & Morgan 1940 *J. Amer. Chem. Soc.* **62**, 16.
White & Bishop 1940 *J. Amer. Chem. Soc.* **62**, 8.

The structure of polychloroprene

By C. J. Birkett Clews

Queen Mary College (University of London), and the Crystallographic Laboratory, Cambridge

(*Communicated by H. R. Robinson, F.R.S.—Received* 12 *June* 1941)

[Plate 1]

Specimens of polychloroprene before and after light treatment have been examined by X-ray methods. There is no change in the crystalline structure, although there are differences in the physical properties ascribed to cross-linking of the long-chain molecules. The unit cell is possibly orthorhombic: $a = 8{\cdot}90$ A, $b = 4{\cdot}70$ A, $c = 12{\cdot}21$ A, and contains four chloroprene (C_4H_5Cl) units.

1. Introduction

In the last few years much attention has been given to the production of highly polymerized substances for use as plastics or rubber substitutes. A large-scale attempt to manufacture synthetic rubber-like materials was made in Germany, 1914–18, owing to a serious shortage of the natural product. The synthetic product, known as methyl rubber, was a polymer of dimethyl butadiene; it was very inferior to ordinary rubber. Since then a number of these polymers have been developed. The most notable are the Buna rubbers of I.G. Farbenindustrie (straight chain polymers of butadiene and interpolymers with butadiene), several Russian butadiene polymers known as Sovprene, and polychloroprene of Du Pont Nemour known as Duprene or Neoprene, developed commercially on account of its superiority over natural rubber for many purposes. The chief disadvantage of most of these materials is their high cost. Many such rubber-like materials have been made from hydrocarbons with a butadiene nucleus, but only in the case of halogen substituted butadienes have the results been completely satisfactory. The most important example is the polymer of 2-chlorobutadiene or chloroprene.

2. Properties of chloroprene and polychloroprene

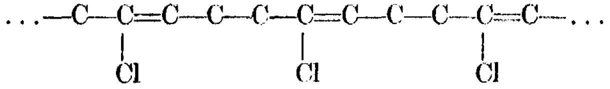
Chloroprene is a colourless liquid boiling at 59·4° C and of density 0·9583 g./c.c. Its chemical formula is

$$CH_2{=}CCl{-}CH{=}CH_2.$$

The exact molecular structure is not known, although an attempt has been made to investigate it by the application of X-ray methods at low temperatures, so far without success (Clews 1938). It is hoped to continue this work on the monomer in the solid state.

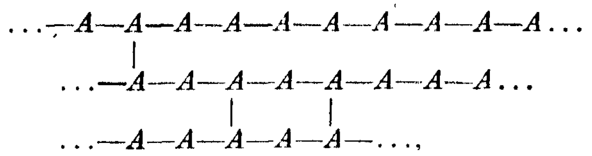
Chloroprene on polymerization may form a number of substances with widely differing properties. The four principal polymers are designated α, β, μ, and ω. The β-polymer is an oily dimer and ω is a granular substance. The α-polymer, which is formed by spontaneous polymerization at normal temperatures, the formation being accelerated by the action of light, corresponds to milled natural rubber. μ-polychloroprene corresponds to vulcanized rubber, and is produced by continuing the reaction which gives the α form. The conversion from the α to the μ form may be carried out by heating the material at 130° C for about 5 min.

Carothers *et al.* (1931) suggested that in the α-polymer the chloroprene molecules are joined to produce a long-chain structure:



$$\ldots-\mathrm{C}-\underset{\mathrm{Cl}}{\underset{|}{\mathrm{C}}}=\mathrm{C}-\mathrm{C}-\mathrm{C}-\underset{\mathrm{Cl}}{\underset{|}{\mathrm{C}}}=\mathrm{C}-\mathrm{C}-\mathrm{C}-\underset{\mathrm{Cl}}{\underset{|}{\mathrm{C}}}=\mathrm{C}-\ldots$$

If the polymer is heated, cross-linkages between the long chains make their appearance, giving



where *A* is a chloroprene unit. It is supposed that this represents the μ-polymer. The exact nature of the cross-linking is somewhat obscure.

The chief advantages of polychloroprene over natural rubber are (1) its better resistance to the action of oils and solvents, (2) its greater resistance to heat, (3) its greater resistance to the effects of oxygen and ozone, and (4) its better ageing properties. The most important property of synthetic rubbers is their greater resistance to oils and some solvents.

Experiments on the action of light on neoprene have been carried out by Melville & Davies (1940), who very kindly supplied the author with information about these experiments and with neoprene specimens for X-ray investigation.

A strip of uncompounded neoprene, contained in an evacuated, water-cooled pyrex vessel, was exposed for 2 hr. to a 25 amp. (200 V) carbon-arc lamp, situated 6 cm. from the surface of the strip. The temperature at the surface was about 40° C. Before illumination, the neoprene was very readily soluble in toluene; viscosity measurements on such a solution indicated an average chain length of about 300 chloroprene units (a molecular weight of 26,000). After illumination, the surface layer, which had been exposed to the light ($\lambda > 3000$ A), was insoluble in toluene; viscosity measurements on the soluble portion revealed no appreciable change in the average chain length. The resistance to stretching of the treated material was much greater than that shown by the normal neoprene.

3. X-RAY EXAMINATION OF POLYCHLOROPRENE

As in the case of natural rubber, an X-ray photograph of unstretched neoprene exhibits a broad diffuse ring. On stretching, the diagram shows a number of discrete interference spots indicating a crystalline or pseudo-crystalline structure, while the diffuse ring, although still present, is less intense.

The specimen under investigation was stretched to its maximum extension and retained in this state by means of a small brass specimen holder which could be mounted in a cylindrical camera of 3 cm. radius. The substance was in the form either of a thin fibre or a strip; in the latter case, the X-ray beam was incident perpendicularly to the plane of the film. In all cases, the fibre axis (or direction of stretching) was parallel to the axis of the camera. The X-ray source was a Philips Metallix tube with a copper anticathode operating at about 35 kV and 18 mA. To obtain satisfactory photographs, it was necessary to make exposures of 24–30 hr. (Ilfex film).

One of the X-ray photographs is reproduced in figure 1*a* (plate 1). It will be noticed that there are quite sharp lines on the equator. It is possible to measure twelve of these and compute the corresponding spacings. The first layer lines may also be clearly seen, although there is a certain diffuseness which prevents the determination of any of the spacings except that from the layer line separation. This corresponds to the identity period along the fibre axis. This is clearly one of the axes of the unit cell of 'crystalline' polychloroprene. It is found to be $4{\cdot}70 \pm 0{\cdot}03$ A, being the mean of a number of measurements on several photographs. This is slightly lower than the value 4·8 A quoted by some other authors (Clark 1932; Katz 1936; Fuller 1940). No particulars of their experiments are given.

The first two equatorial lines are due to reflexion of Cu $K\beta$ and $K\alpha$ from the same set of planes, assumed to be the second-order reflexions from the (100) planes. The next reflexion on the equator is also fairly intense and probably corresponds to the (001) planes, possibly the second or third order. The various possibilities will be discussed later.

Photographs were also taken with specimens of the light-treated polychloroprene in the stretched state. One of these diagrams is reproduced for comparison (figure 1*b*, plate 1). Measurement shows that the spacings calculated from the layer line separation and the equatorial lines are identical with those for the untreated material. There are, however, two differences between the two photographs:

(1) The equatorial interferences are sharper on the photograph of the treated material.

(2) There is evidence of complete Debye-Scherrer rings, their intensity being much sharper where they cut the equator.

4. DISCUSSION OF THE STRUCTURE OF POLYCHLOROPRENE

It is at once evident from the X-ray photographs and the measurements that there is no difference in the crystalline 'fine structure' of the two specimens; further, it is obvious that in both cases there is an orientation of the long chains so that most of them lie parallel to the fibre axis, caused by stretching. In the previous section, it was mentioned that the treated material was insoluble in toluene, and also that it showed increased toughness. Both these facts suggest cross-linking of the long chains induced in the substance by illumination. It is estimated that, with an average chain length of 300 units, a 1 % degree of cross-linking would be sufficient to produce insolubility (Melville & Davies 1940). This is supported by the X-ray results; the increased sharpness of the equatorial interferences after the illumination of the specimen may well be due to cross-linking. The same phenomenon is observed when natural rubber is vulcanized. It is thought that during vulcanization sulphur bridges are formed within a given chain and linkages from one chain to another.

The complete Debye-Scherrer rings, with the much greater intensity on the equatorial axis, indicate that there is a marked orientation of the crystalline regions, so that the polymerized chains lie parallel to the fibre axis, but that some of the crystalline regions are orientated at random. It is possible that the diffuseness of the layer lines with both specimens is due to a slight variation in the periodicity along the fibre axis. Katz (1936) has also reported a 'fuzziness' of the layer lines.

The theory of the 'crystallite' or 'micellar' structure of high polymers, and in particular rubber, has been discussed elsewhere (Clews & Schoszberger 1938), so it will not be considered now except to see how far it may be applied to the present problem. In ordinary rubber there is no evidence of cross-linking to form a net structure, so that there is practically complete ordering of the crystallites on stretching; the long main valence chains are all parallel to the stretching direction, and with untreated polychloroprene there is also fairly complete orientation. It is very probable that cross-linking in the light-treated material prevents the complete ordering of the chains parallel to the fibre axis, but that short lengths of several chains may be parallel to one another and not parallel to the fibre axis. These latter would give rise to the random orientated crystallites. The idea of cross-linking is also supported by the decreased extensibility after treatment.

Let us now consider the crystalline structure. The periodicity along the fibre direction or b-axis is 4·70 A, and assuming the first three equatorial reflexions to be $(200)\,\beta$, $(200)\,\alpha$ and (003), respectively, the unit cell has dimensions

$$a = 8{\cdot}90\ \text{A}; \quad b = 4{\cdot}70\ \text{A}; \quad c = 12{\cdot}21\ \text{A}.$$

(It is assumed $\alpha = \beta = \gamma = 90^\circ$.) There is the possibility that the third line may be the second and not the third order of (001), that is the unit cell, again assumed orthorhombic, is

$$a = 8{\cdot}90\ \text{A}; \quad b = 4{\cdot}70\ \text{A}; \quad c = 8{\cdot}14\ \text{A}.$$

The first unit cell has a volume 511 A^3; substituting in the expression for the density

$$\rho = 1{\cdot}65\, nM/V,$$

and calculating n, the number of chloroprene units in the cell, we obtain the value $n = 3{\cdot}80$. (The experimentally determined density is 1·086 g./c.c., and the molecular weight, M, of the chloroprene unit is 88·5.) n must, however, be a whole number; obviously, $n = 4$, giving a calculated density 1·14 g./c.c. Taking the second set of cell dimensions and four molecules per unit cell, the density is 1·72 g./c.c. It is reasonable to suppose that the density will be slightly greater when the specimen is stretched, but the density calculated from the second set of cell data is much higher than can be accounted for by stretching. (The density of natural rubber is 0·906 g./c.c. and 0·956 g./c.c. when stretched.) It would be necessary to assume three chloroprene units in the cell, and this is incompatible with orthorhombic symmetry. An orthorhombic cell,

$$a = 9{\cdot}0\ \text{A}; \quad b = 4{\cdot}79\ \text{A}; \quad c = 8{\cdot}23\ \text{A},$$

has been suggested by Krylov (1935) from electron diffraction measurements. His results are reported briefly by Fuller (1940), but the number of units in the cell and the density are not mentioned. It appears that the

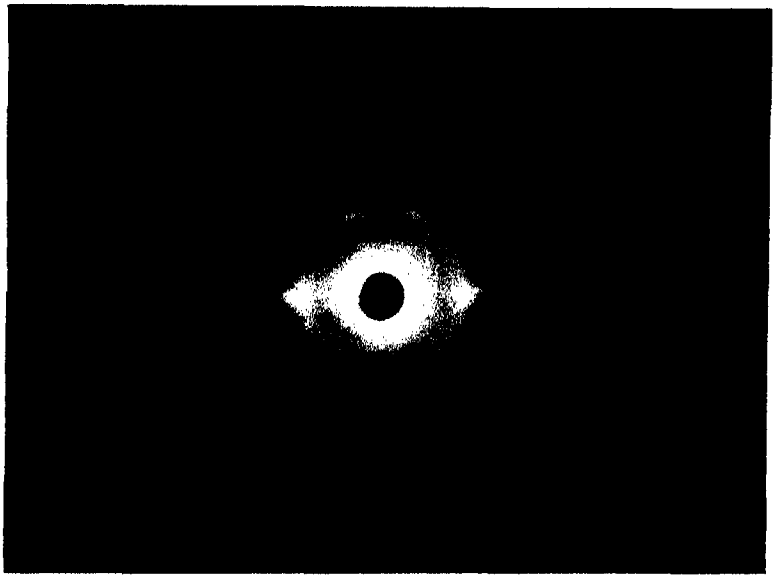


FIGURE 1*a*. Untreated polychloroprene

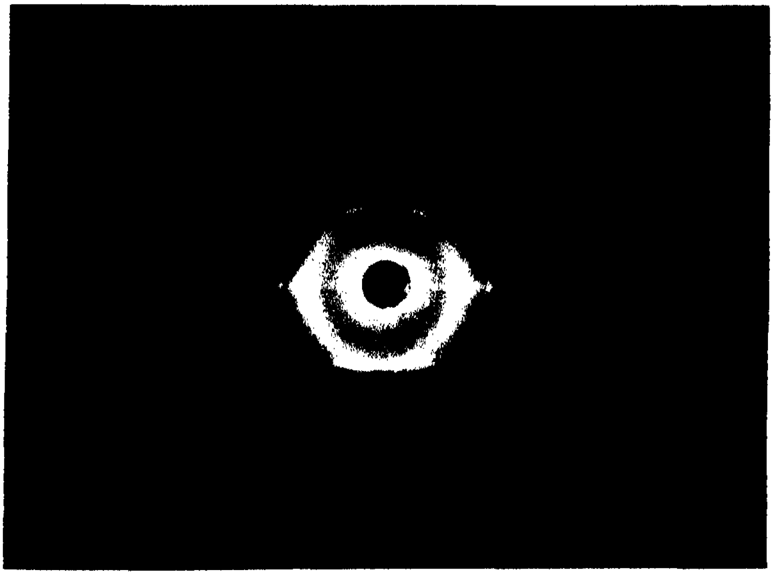


FIGURE 1*b*. Light-treated polychlorene.

measured density and this smaller cell cannot be reconciled. It may therefore be supposed that the first cell given is probably the more correct, although this can be no more than a rough approximation to the true one. There are four chloroprene groups associated with polychloroprene chains passing through the unit cell and lying, on the whole, parallel to the *b*-axis. From the repeat distance, 4·70 A, it is evident that the chains are in the *trans* form. Even on this assumption there must be some modification of either the bond angles or bond lengths from the usual values, or it must be assumed that the molecule is not planar. A similar difficulty was found in fitting the isoprene groups into the rubber unit cell.

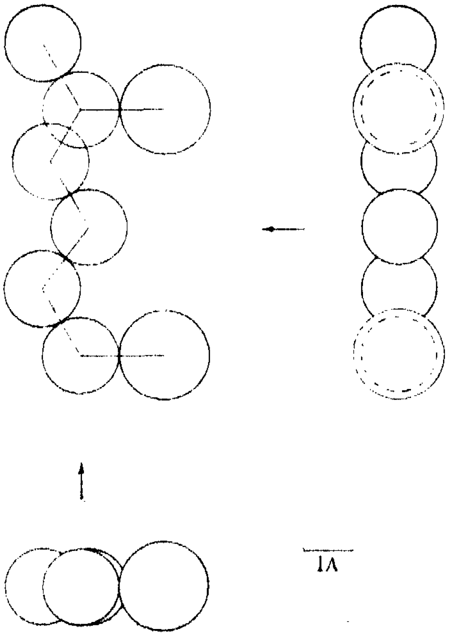


FIGURE 2

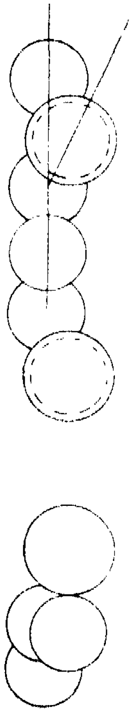


FIGURE 3

Part of a planar chloroprene chain is illustrated in figure 2. The single-bond distance is taken as 1·5 A, the double-bond distance as 1·2 A, and the bond angles are 109° and 120°. The radius of the chlorine atom is 1·8 A. The repeat distance is 4·9 A. A possible non-planar molecule is shown in figure 3. This has a repeat distance of 4·7 A in accordance with the X-ray data. The hydrogen atoms are not represented in these figures.

From the dimensions of the molecule and of the unit cell, it is obvious that the chloroprene groups must be packed closely together, possibly with intermolecular distances less than those usually expected.

In an endeavour to index the other equatorial reflexions, it was soon realized that this would not be possible if all sets of indices having $k \neq 0$ were excluded. This is not unreasonable because evidently we are not dealing with a 'single crystal rotation' type of diagram produced by rubber, but with a Debye-Scherrer photograph showing preferred orientation. The measured spacings and their corresponding indices and visually estimated intensities are given in table 1. The indexing can have no great significance, there being far too few spacings either to confirm or refute the correctness of the cell chosen.

TABLE 1

Camera radius 3 cm. Copper $K\alpha\beta$ radiations

number of line	$\sin^2\theta$	d in A	hkl	intensity
1	0·20682	1·69	323	*w*
2	0·18243	1·80	413	*ww*
3	0·17000	1·87	024	*w*
4	0·12093	2·22	400	*w*
5	0·09635	2·48	311	*w*
6	0·08927	2·58	—	*ww*
7	0·08114	2·71	—	*mw*
8	0·06115	3·12	211	*ms*
9	0·05266	3·36	204	*s*
10	0·03587	4·07	003	*s*
11	0·03016	4·45	200α	*ss*
12	0·02455	4·45	200β	*s*

5. CONCLUSION

Specimens of polychloroprene and light-treated polychloroprene have been examined by X-ray methods. There is evidently no change in the crystalline structure of the two substances, although there are differences in physical properties which are ascribed to cross-linking of the long polychloroprene chains. This conclusion is confirmed by the X-ray work. The unit cell of polychloroprene is probably orthorhombic with dimensions $a = 8{\cdot}90$ A, $b = 4{\cdot}70$ A, $c = 12{\cdot}21$ A. From the experimentally determined density of 1·086 g./c.c. it is deduced that there are four chloroprene units in the cell. The X-ray density (for the stretched material) is 1·14 g./c.c. The long chains lie parallel to the b-axis (the direction of stretching), four of these chains passing through the unit cell. It is concluded, from the identity period, that the chain is not planar and a possible spatial configuration is given.

The experimental work has been carried out in the Crystallographic Laboratory in Cambridge, and I am greatly indebted to Professor Sir Lawrence Bragg, F.R.S., and to Dr A. J. Bradley, F.R.S., for their generous hospitality. The Rubber Manufacturers' Research Association, Dr W. J. S. Naunton of Imperial Chemical Industries, and Professor H. W. Melville, F.R.S., kindly supplied me with specimens. The work has been assisted by the Rubber Growers' Research Board and a grant from the Central Research Fund of the University of London. Finally, my best thanks are due to Professor H. R. Robinson, F.R.S., for his interest and encouragement.

REFERENCES

Carothers, Williams, Collins & Kirby 1931 *J. Amer. Chem. Soc.* **53**, 4203.
Clark, G. L. 1932 *Applied X-rays*, p. 450.
Clews, C. J. B. 1938 *Nature, Lond.*, **141**, 513.
Clews, C. J. B. & Schoszberger, F. 1938 *Proc. Roy. Soc.* A, **164**, 491.
Fuller, C. S. 1940 *Chem. Rev.* **26**, 143.
Katz, J. R. 1936 *Trans. Faraday Soc.* **32**, 91.
Krylov 1935 *J. Exp. Theor. Phys. U.S.S.R.* **5**, 524.
Melville & Davies 1940 Private communication.

The lattice spacings of the primary solid solutions in magnesium of the metals of Group IIIB and of tin and lead

BY GEOFFREY VINCENT RAYNOR

(*Communicated by W. Hume-Rothery, F.R.S.—Received 24 June* 1941.—*Revised* 11 *September* 1941)

The lattice spacings of the primary solid solutions in magnesium of aluminium, gallium, thallium, tin and lead have been accurately measured. The new data confirm that the effect of increasing the valency of the solute is to tend to expand the lattice of magnesium. Thus, although the interatomic distance for tin is smaller than the corresponding distances for indium and thallium, the contraction in the '*a*'-spacing of magnesium-tin alloys is markedly less than in the indium and thallium alloys. Lead expands the lattice of magnesium.

The '*a*'-spacing-composition curves for the systems studied are smooth and continuous. In contrast, the '*c*'-spacing curves show a change in direction at an electron concentration of approximately 2·0075, that is, at 0·75 atomic per cent of a three-valent solute and 0·375 of a four-valent solute. It is at this point that an overlap of electrons from the first Brillouin Zone for the magnesium structure occurs in the 'c' direction, in addition to the already existing overlap in the 'a' direction. For solutes of favourable 'size factor' the variation in axial ratio with composition, after the onset of the second overlap, is proportional to the electron concentration, and is therefore twice as great per atomic per cent for a four-valent solute as for a three-valent solute.

The apparent atomic diameters of the above elements in dilute solid solution in magnesium, as deduced by extrapolation of the lattice spacing-composition curves to 100 % of solute, are in agreement with the requirements of the Brillouin Zone theories of metals. The results confirm the suggestion that indium, thallium, white tin and lead are incompletely ionized in the elementary state.

1. INTRODUCTION

A systematic investigation of the primary solid solutions of many elements in copper and silver (Hume-Rothery, Mabbott & Channel-Evans 1934) established the principal factors affecting solid solution formation. It was concluded that the predominant factors were the 'size-factor' relationships and the valencies of the solvent and solute elements. Where the closest distance of approach of the atoms in the crystals of the elements differed by more than 14 % of that of the solvent, the primary solid solution was very restricted, and the size factor was said to be unfavourable. Where the closest distances of approach were within this limit, wide solid solutions were formed and the size factor was said to be favourable. This empirical rule enabled the phenomena of alloy formation to be interpreted, to a first approximation, in terms of constants of the constituent elements. A later examination of the lattice distortion produced in silver and copper by elements of favourable size factor with respect to these solvents showed that clear numerical relations existed.

In a paper discussing magnesium alloys in the light of the theory of copper and silver alloys (Hume-Rothery & Raynor 1938), it was shown that the same considerations applied to solid solution formation in magnesium, provided that allowance was made for the highly electropositive nature of this metal. The lattice distortions produced in magnesium by indium, cadmium and silver were reported in a previous paper (Raynor 1940). Equal atomic percentages of these elements contract the 'a'-spacing of magnesium by amounts proportional to 1:2:5 respectively. It was pointed out later (Hume-Rothery & Raynor 1940) that although the a-spacing-composition curves for the systems magnesium-indium and magnesium-aluminium are smooth and continuous, the c-spacing-composition curves show an abrupt change in direction at about 0·75 atomic % of the solute.

The present paper represents an extension of this work. The mean lattice spacings of the primary solid solutions of gallium, thallium, tin and lead in magnesium have been measured and further experiments made on the magnesium-aluminium alloys. According to the empirical size-factor rule, the size factors of indium and thallium are favourable, that of aluminium is on the borderline of the favourable zone, while that of gallium is unfavourable with respect to magnesium. Tin and lead have favourable size factors, but their primary solid solutions in magnesium are restricted by the effect of stable compound formation.

Throughout the paper the term 'mean lattice distortion' is used to denote the differences between the lattice spacings of the solid solutions and the solvent metal, as measured by the ordinary powder methods of X-ray analysis.

2. GENERAL EXPERIMENTAL PROCEDURE

The experimental technique employed was similar to that fully described in a previous paper (Raynor 1940). The alloys were prepared by melting the constituents together in alumina crucibles, and casting into copper moulds. After annealing the ingots to equilibrium, filings were prepared in an argon atmosphere, annealed in silica tubes to relieve strain at temperatures within the homogeneous area of the equilibrium diagram, and quenched in water. Fine filings separated by sieving were mounted on a hair for exposure and the unsieved residue sealed in vacuo for analysis.

The compositions of all alloys were established by chemical analysis, both constituents being determined throughout as a check on contamination. Tests showed that the compositions by analysis of the sieved filings and the unsieved residues were identical, so that in general the latter were used for analytical work. The figures given in the Appendix are therefore based on the analysis of the filings after all annealing treatments had been completed, that is, in the condition in which they were exposed to X-rays.

Lattice constants were determined in the usual manner, using a cylindrical camera of diameter 19 cm., and copper K_α radiation from a Metropolitan-Vickers demountable X-ray tube. The temperature of the camera during exposure was kept constant and was measured by means of a thermocouple. The lattice spacings quoted in this paper have been reduced to a common temperature of 20° C, assuming for this purpose that the coefficients of expansion of the dilute alloys were substantially the same as those of magnesium.

The camera was calibrated against quartz, as recommended by Bradley & Jay (1933), whose $\cos^2\theta$ extrapolation method was used throughout in order to eliminate errors due to absorption, film shrinkage, eccentricity of specimen, etc. In the course of the work, Dr H. Lipson kindly informed the author that extensive work at the Crystallography Department, University of Cambridge, had shown that the constants recommended by Bradley & Jay gave the standard angle of the camera, as defined by shadows of fixed knife edges, consistently 0·04° too high. The present results were therefore calculated using a standard angle of 84·26° instead of the value 84·30° originally assumed.

Duplicate measurements of the same film gave results agreeing within 0·0001 A, while duplicate photographs of the same alloy gave results also agreeing within this limit after correction for the temperature of the exposure. It is estimated that the values of the a-spacings are accurate to within ± 0·0001 A, while the values of the axial ratios and c-spacings are estimated to be accurate to within ± 0·0001 and ± 0·0003 A respectively.

3. MATERIALS USED

The magnesium used was presented by the National Physical Laboratory and was 99·95 % pure; the chief impurities were iron (0·03 %), silicon (0·01 %) and aluminium (0·01 %). Particulars of the solute metals used are as follows:

Aluminium: the purity exceeded 99·99 %, the impurities being iron (0·0008 %), silver (0·0028 %) and copper (0·001 %).

Gallium: supplied by the Kaliwerke, Aschersleben, Germany, and refined before use by Messrs Johnson, Matthey and Co., Ltd.

Thallium: the metal was kindly presented by the Goldsmiths' Metallurgical Laboratory, Cambridge, and was 99·99 % pure.

Tin: 'Chempur' tin, as supplied by Messrs Hopkins and Williams, was employed.

Lead: this metal was kindly presented by the Cookson Lead and Antimony Company, Ltd., and was spectroscopically pure.

4. EXPERIMENTAL RESULTS

(a) *Pure magnesium*

Five separate determinations of the lattice spacings of magnesium gave the axial ratio at 20° C as 1·6236 (assuming θ_k for camera = 84·26). The values of the a-spacing varied between 3·20255 and 3·20265 A, with a mean value of 3·20259 A, while the mean value of the c-spacing was 5·19973 A. These values are in excellent agreement with the results of Jette & Foote

(1935), who reported an a value of 3·20300 A at 25° C (= 3·20258 A at 20° C) with c/a equal to 1·62355. They are also in good agreement with a recent measurement of Ievins, Straumanis & Karlsons (1938), which gave $a = 3{\cdot}20280$ A at 25° C (= 3·20238 A at 20° C) with c/a equal to 1·62353.

(*b*) *The alloy systems*

The results obtained for the lattice-spacing measurements with alloy specimens are tabulated in the Appendix to this paper.

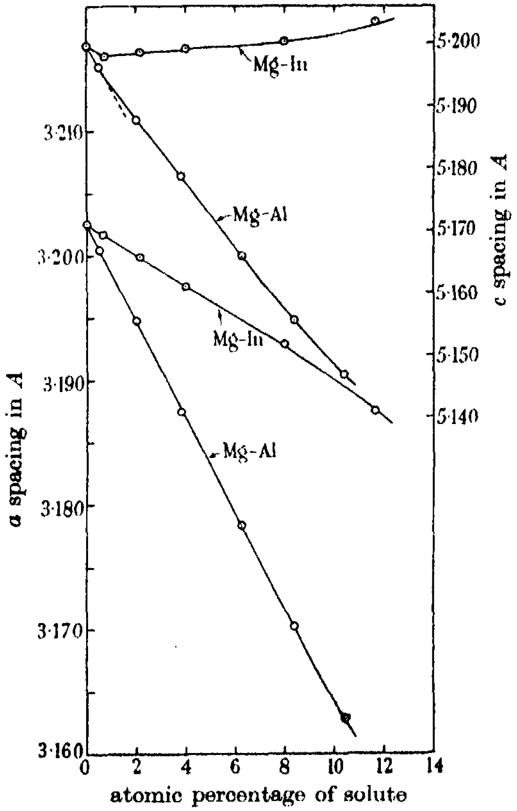


FIGURE 1

Figure 1 shows the lattice-spacing-composition curves for the primary solid solution of aluminium in magnesium, together with those for the magnesium-indium system for comparison. The experimental points for the a-spacings lie on a smooth curve from which no point deviates by more than 0·0002 A. The curve is initially linear; above the approximate limit of 6 atomic % the contraction is slightly less than would be expected from a linear law.

The spacing-composition curves for the systems magnesium-gallium and magnesium-thallium are shown in figure 2 (*a*-spacings) and figure 3 (*c*-spacings). The *a*-spacing curve for the gallium alloys falls sharply from the value for magnesium and is initially linear. In the region of 2 atomic % of gallium the curve, from which no point differs by more than 0·00015 A, departs from linearity in the same sense as in the aluminium alloys. For the magnesium-thallium alloys the *a*-spacing curve, from which no point differs by more than 0·0001 A, again appears to be linear for the dilute alloys; the contraction per atomic per cent later becomes more marked, in contrast to the alloys with aluminium and gallium. Similar behaviour is shown by the magnesium-indium *a*-spacing curve.

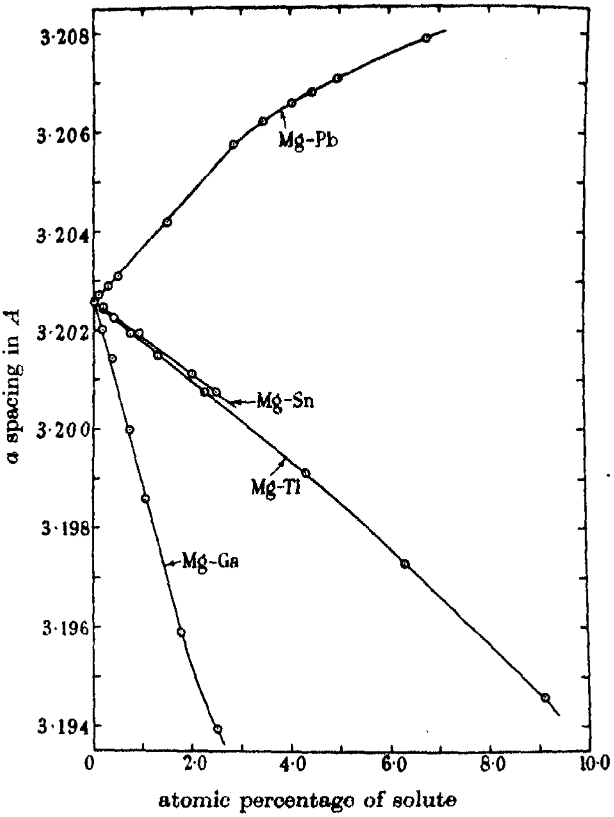


FIGURE 2

Whereas the a-spacing curves for these alloy systems are smooth and continuous, the curves for the c-spacings all show an abrupt change of direction at a composition of 0·75 atomic % of solute. The contraction per atomic per cent in the aluminium and gallium alloys is less above this limit than below, while the initial contraction is followed by an expansion in the case of the alloys with indium and thallium.

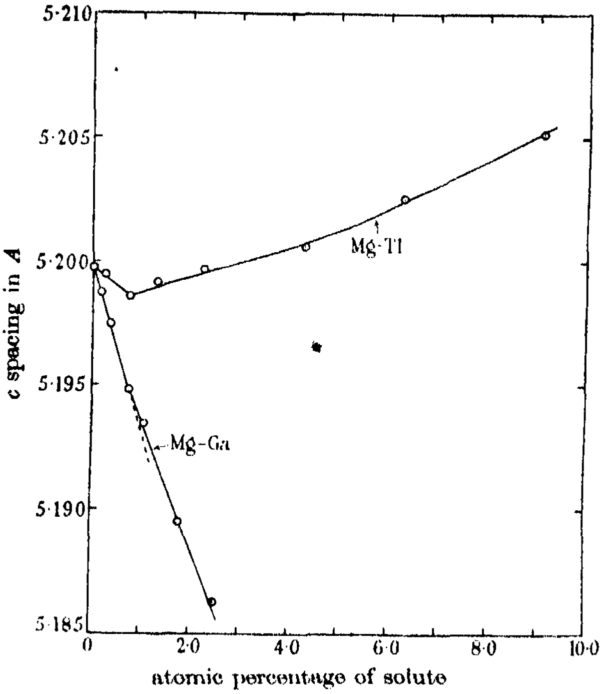


FIGURE 3

Figures 2 and 4 show respectively the a-spacing curves and the c-spacing curves for the primary solid solutions of tin and lead in magnesium. For the magnesium-tin system, a straight line, from which no point deviates by more than 0·00005 A, may be drawn through the determined a-spacings. The lattice distortion is less, per atomic per cent, than in the case of the indium and thallium alloys. The c-spacing curve is of the same general type as those for the magnesium-indium and magnesium-thallium systems, showing an abrupt change of direction. Since this effect occurs at such a small tin content, it is difficult to determine accurately the exact composition at which it takes place. It must, however, lie between 0·2 and 0·4 atomic % of tin.

It is seen from figure 2 that the magnesium-lead system, in contrast to the other systems investigated, shows an expansion of the a-spacing. The a-spacing curve, from which no point differs by more than 0·0001 A, appears again to be initially linear. In the region of 3 atomic % of lead the linearity is lost and the expansion per atomic per cent becomes progressively less. The experimental points suggest that the change is gradual rather than abrupt. The form of this curve was confirmed by an investigation published by Foote & Jette (1940) after the present work had been completed. The a-spacings reported by these authors are scattered by amounts of the order of 0·0003 A about a smooth curve which, after reduction to 20° C, lies below that of figure 2 but differs from it at no point by more than 0·0002 A. This agreement is very satisfactory. According to these authors, who did not examine alloys containing less than 0·51 atomic

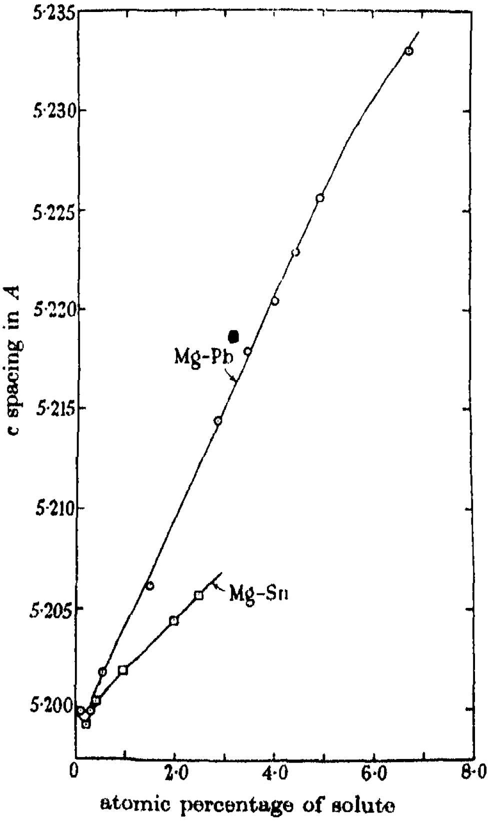


FIGURE 4

% of lead, the c-spacing curve is also smooth and continuous. Figure 4, however, shows that there is again a change of direction, the expansion in the very dilute alloys being much less marked than that above a critical limit. The exact composition of this limit is again difficult to locate, but must lie between 0·3 and 0·5 atomic % of lead. The experimental points of Foote & Jette are not inconsistent with this change in direction.

5. DISCUSSION

It has been stated above that the size-factor conception of Hume-Rothery and his co-workers enabled alloy formation to be interpreted in terms of the atomic diameters of the constituent metals, which were assumed to be given by the closest distances of approach of the atoms in the crystals of the elements. On this basis the size factors for gallium, indium, thallium, and lead with respect to magnesium are difficult to assess, since the interatomic distances in metallic indium, thallium and lead are affected by incomplete ionization (Hume-Rothery & Raynor 1938), while gallium has a complex crystal structure such that there are several close distances of approach between atoms. An estimate of the atomic diameters in structures showing incomplete ionization can, however, be made. The difference between the atomic diameters of white tin (incompletely ionized) and grey tin (fully ionized) is of the order of 0·3 A, so that by subtraction of this difference from the observed interatomic distances in metallic indium, thallium and lead, approximate values are obtained, which may be taken as corresponding with complete ionization. By this arbitrary procedure it is estimated that the size factors with respect to magnesium of the solutes studied in the present work are in the following order:

Lead (most favourable), thallium, indium, tin, aluminium (borderline), gallium (unfavourable).

From the curves shown in figures 1 and 2 for the three-valent solutes, it is seen that the distortion produced corresponds qualitatively with this order. The distortion produced by equivalent amounts of gallium and aluminium is markedly greater than that produced by indium and thallium, while the curve for thallium lies slightly above that for indium. The a-spacing curves for dilute solutions of gallium and aluminium in magnesium are similar; the distortion produced by gallium would, however, be expected to be greater than that produced by aluminium if the gallium size factor is estimated from the closest distance of approach of atoms in metallic gallium. This behaviour is referred to below.

A comparison of the a-spacing curves for the primary solid solutions of indium, cadmium and silver in magnesium (Raynor 1940) showed that increasing the valency of the solute decreased the lattice distortion in the a direction, this effect being due to an increase in the overlap of electrons in the a direction from the first Brillouin Zone for the magnesium structure as the electron concentration increases. Figure 2 shows that this principle holds also for four-valent solutes. An increase of 1 atomic % of a four-valent solute gives an increase of 0·02 overlapping electrons per atom, as compared with an increase of 0·01 per atom for 1 atomic % of a three-valent solute. The relative expansion of the a-spacing is therefore greater, per atomic per cent, for four-valent solutes. In agreement with this the magnesium-tin distortion curve lies above those for the thallium and indium alloys, although the arbitrary size factor is less favourable.

It is of interest to note that, according to the present work, the volume of the unit cell of magnesium-tin alloys remains constant, within the limits of experimental error, above the composition corresponding to the direction change in the c-spacing curves.

In dilute solution, equiatomic percentages of indium, cadmium and silver contract the a-spacing of magnesium by amounts proportional to 1:2:5 respectively. The distortion produced by tin, which is in the same period of the periodic table as indium, cadmium and silver, bears no simple relation to that produced by these elements. This is not unexpected, since the relatively high electrochemical effect in the magnesium-tin system tends to contract the lattice and obscure purely electronic effects.

Taking the present results, and combining them with previous results for indium, cadmium and silver, it is found that equiatomic percentages of these elements contract the a-spacing of magnesium by amounts proportional to 1:1·75:3·5:8·75, while the volume of the unit cell, for compositions lying above that corresponding to the direction change, is decreased by amounts proportional to 0:1:3:9. It is curious that both series of numerical factors may be represented by empirical expressions of the same form. The expression $2F_a + 3 = 20/V$, where F_a is the numerical factor for the a-spacing curves and V the valency of the solute, leads to the values 1, 1·8, 3·5, and 8·5 for F_a, while the expression $2F_v + 6 = 24/V$ gives the values 0, 1, 3, and 9 for the volume factor F_v.

Although the a-spacing curves for the solid solutions of aluminium and indium are smooth and continuous, the c-spacing curves show a change in direction at 0·75 atomic % of solute (figure 1). This behaviour was noted by Hume-Rothery & Raynor (1940), and was taken to imply that, although in metallic magnesium with two electrons per atom the overlap of the first Brillouin Zone is in the *a* direction only, the structure is so near the stage at which an overlap in the *c* direction sets in that the addition of less than one electron per hundred atoms causes the *c* overlap to take place. Once this overlap is established, the stresses produced by the overlapping electrons tend to expand the lattice in the *c* direction; this leads to a change in direction of the *c*-spacing curve at the critical composition. Figure 3 shows clearly that the magnesium-gallium and magnesium-thallium systems exhibit the same behaviour. The effect is very clearly shown by the gallium alloys, in which the first small additions of gallium lower the axial ratio to 1·6234 at 0·75 atomic %, while further additions increase it. The resulting change in direction of the *c*-spacing curve is slight but is outside the limits of experimental error.

If the above interpretation of the effect is correct, similar behaviour should be observed in the *c*-spacing curves of magnesium alloys with four-valent elements at a composition of 0·375 atomic %, which corresponds with an electron concentration of 2·0075. Figure 4 shows that this prediction is fulfilled for the solid solutions in magnesium of tin and lead, which are the only four-valent solutes having a sufficiently wide solid solution range in magnesium for the effect to be tested.

Confirmation of the general interpretation is given by the observation that, for solutes of favourable size factor, the increase of axial ratio per atomic per cent above the critical limit is approximately twice as great for four-valent solutes as for three-valent solutes. The axial ratio changes therefore depend directly upon the electron concentration, or the number of overlapping electrons per atom. Figure 5 shows the axial ratios for the indium, thallium, tin and lead alloys plotted against composition. In figure 6*a* the axial ratios for the alloys with indium and tin, which both belong to the same period, are plotted against electron concentration, while figure 6*b* shows the corresponding plot for the thallium and lead alloys. In each case the points fall on a single curve in dilute solution.

The results have been discussed in terms of the empirical size-factor relationships, with arbitrary correction where necessary for incomplete ionization of the solute. It was pointed out, however, by Hume-Rothery & Raynor (1940) that the interatomic distances in crystals of alloys cannot be accounted for by assigning a fixed radius to each kind of atom. If the alloy structure corresponds with an overlap in any direction from the first Brillouin Zone, this direction in the lattice will be relatively expanded, so that the apparent atomic diameter of the solute as deduced by extrapolation of the lattice spacing curve to 100 % of solute is greater than in the absence of overlap. The apparent atomic diameter of aluminium in

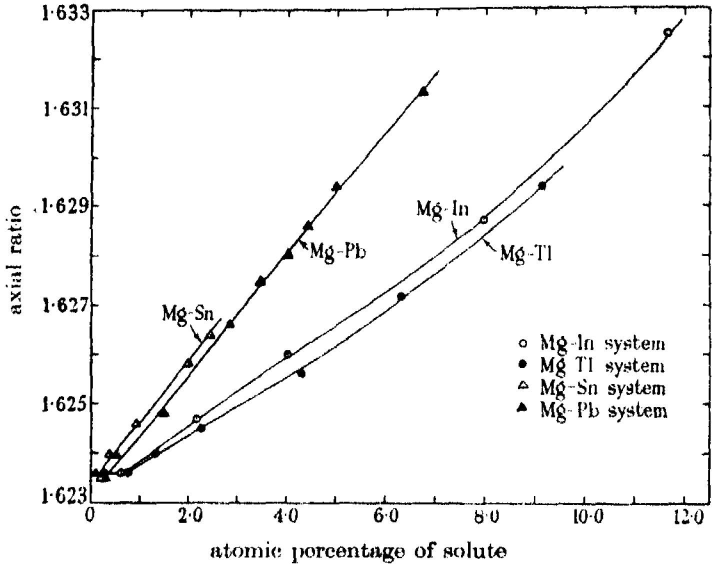


Figure 5

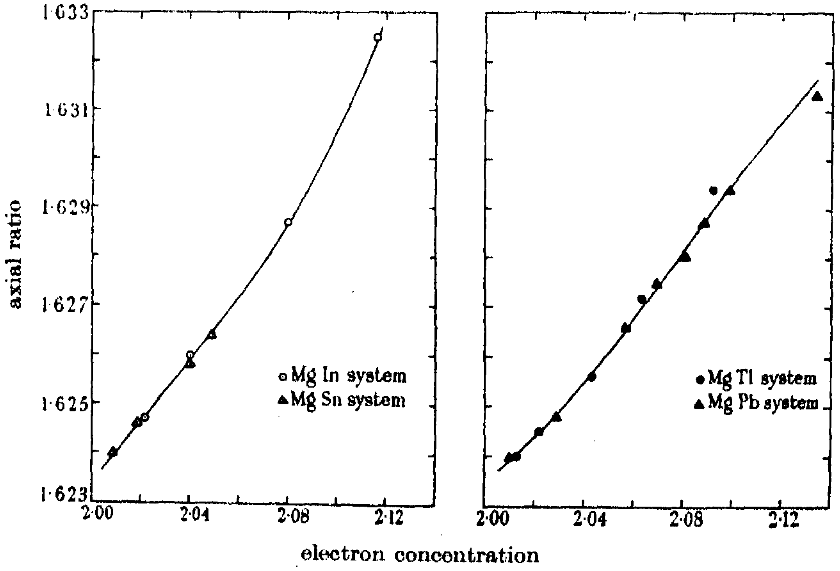


Figure 6

solution in copper, silver or gold where there is no overlap is, for example, less than the value obtained by extrapolating the a-spacings for dilute magnesium-aluminium alloys, for which the Brillouin Zone is overlapped in the a direction. The value obtained in the latter case is close to the normal atomic diameter in metallic aluminium, in agreement with the fact that both involve zone overlaps.

The apparent diameter of the indium atom obtained by extrapolating the a-spacings of magnesium-indium alloys is 3·09 A, as compared with a closest distance of approach of 3·24 A in the crystal of the metal. This behaviour is ascribed to the fact that, in addition to the electrochemical effect being more marked for these alloys than for magnesium-aluminium alloys, metallic indium shows some tendency towards incomplete ionization.

In the same way, the apparent atomic diameters of thallium and lead obtained by extrapolating the a-spacing curves have the values 3·12 and 3·31 A respectively, as compared with closest distances of approach 3·40 and 3·49 A in the metals. Since both the alloys and the metals involve electron overlaps, this may be taken as confirmation of incomplete ionization in the metallic state. It is interesting to note that the difference between the 'extrapolated' and 'metallic' atomic diameters is very close to the arbitrary correction previously used to correct the observed metallic values for incomplete ionization.

In the case of tin, the atomic diameter deduced by extrapolation of the a-spacing curve for magnesium-tin alloys is 3·13 A, and corresponds to a hypothetical state of the metal with coordination number 12, whereas white tin has a complex crystal structure with two sets of close neighbours. In view of the uncertainty of corrections for changes in coordination number, it is not possible to compare the extrapolated and metallic values. If, however, the value 3·13 A is corrected for a change in coordination number from 12 to 4 according to the Goldschmidt rule, the result (2·77 A) is close to the interatomic distance in grey tin (2·79 A) which has the coordination number 4.

The crystal structure of gallium is complex, and results in the existence of four sets of neighbours, at 2·437, 2·706, 2·736, and 2·795 A respectively, for each atom. The apparent atomic diameter of gallium in dilute solution in magnesium, as obtained by extrapolation of the a-spacings, is 2·829 A, and is thus considerably larger than the closest distance of approach in the metal. For the purpose of assessing size factors as defined above, the atomic diameter of gallium was assumed to be measured by this closest distance of approach, but it was pointed out by Hume-Rothery, Reynolds & Raynor (1940) that this arbitrary procedure gave too low a value to be consistent with the experimental results for copper, silver, and gold alloys with gallium. The present results confirm this for magnesium alloys. The survey of copper, silver and gold alloys suggested that an effective atomic diameter of 2·7–2·8 A for gallium would be consistent with the facts; the value 2·829 A obtained in the present work is close to this.

The second, or d, distance of approach in the close packed hexagonal structure is not directly related to the Brillouin Zone overlaps, since it involves both the a- and the c-spacings. It should, however, be noted that the extrapolated atomic diameter of gallium for the d distance of approach in the solid solution in magnesium *before* the second overlap begins is 2·80 A, and is thus less than for the overlapped a direction, in agreement with the theory. This effect is general for the alloy systems studied.

The lattice-spacing curves for the solid solutions in magnesium of the metals of group IIIB and of tin and lead therefore support the suggestion of incomplete ionization for metallic indium, thallium, tin and lead.

The author wishes to thank Professor C. N. Hinshelwood, F.R.S., for laboratory accommodation and many other facilities, and Dr W. Hume-Rothery, F.R.S., in whose laboratory the work was carried out, for his interest. Grateful acknowledgement is also made to the Department of Scientific and Industrial Research for a Senior Research Award, to the Government Grants Committee of the Royal Society for a grant towards the expenses of analysis, and to the Aeronautical Research Committee and the Imperial Chemical Industries, Ltd., for other assistance.

REFERENCES

Bradley, A. J. & Jay, H. 1933 *Proc. Phys. Soc.* **45**, 507.
Foote, F. & Jette, E. R. 1940 *Tech. Publ. Amer. Inst. Min. Engrs*, No. 1248.
Hume-Rothery, W., Mabbott, G. W. & Channel-Evans, K. M. 1934 *Phil. Trans.* A, **233**, 1.
Hume-Rothery, W., Lewin, G. F. & Reynolds, P. W. 1936 *Proc. Roy. Soc.* A, **157**, 167.
Hume-Rothery, W. & Raynor, G. V. 1938 *J. Inst. Met.* **63**, 227.
Hume-Rothery, W. & Raynor, G. V. 1940 *Proc. Roy. Soc.* A, **177**, 27.
Hume-Rothery, W., Reynolds, P. W. & Raynor, G. V. 1940 *J. Inst. Met.* **66**, 191.
Ievins, A., Straumanis, M. & Karlsons, K. 1938 *Z. Phys. Chem.* [B], **40**, 347.
Jette, E. R. & Foote, F. 1935 *J. Chem. Phys.* **3**, 605.
Jones, H. 1934 *Proc. Roy. Soc.* A, **147**, 400.
Owen, E. A. & Roberts, E. W. 1939 *Phil. Mag.* (7), **27**, 294.
Raynor, G. V. 1940 *Proc. Roy. Soc.* A, **174**, 457.
Raynor, G. V. & Hume-Rothery, W. 1939 *J. Inst. Met.* **65**, 379.

APPENDIX

The following table contains the measured lattice spacings for the alloy systems mentioned in the text. All values have been reduced to 20° C.

system	alloy no.	wt. % solute	at. % solute	'*a*'-spacing in A at 20° C	'*c*'-spacing in A at 20° C	axial ratio *c*/*a*
pure Mg	—	—	—	3·20259	5·19973	1·6236
MgAl	1	0·55	0·50	3·20045	5·19625	1·6236
	2	2·21	2·00	3·19484	5·18778	1·6238
	3	4·18	3·79	3·18755	5·17881	1·6247
	4	6·95	6·31	3·17842	5·16589	1·6253
	5	9·29	8·45	3·17029	5·15553	1·6262
	6	11·44	10·43	3·16282	5·14717	1·6274
MgGa	1	0·48	0·17	3·20201	5·19878	1·6236
	2	1·05	0·37	3·20144	5·19753	1·6235
	2	1·05	0·37	3·20147	5·19758	1·6235
	3	2·11	0·75	3·19998	5·19484	1·6234
	4	2·91	1·04	3·19858	5·19354	1·6237
	5	4·89	1·76	3·19588	5·18947	1·6238
	6	6·90	2·50	3·19396	5·18636	1·6238
MgIn	1	2·88	0·63	3·20165	5·19820	1·6236
	2	9·45	2·16	3·19990	5·19888	1·6247
	3	16·46	4·01	3·19756	5·19923	1·6260
	4	29·04	7·98	3·19287	5·20023	1·6287
	5	38·35	11·65	3·18745	5·20351	1·6325
MgTl	1	2·04	0·25	3·20244	5·19948	1·6236
	2	5·96	0·75	3·20193	5·19865	1·6236
	3	10·13	1·31	3·20150	5·19924	1·6240
	4	16·09	2·23	3·20075	5·19962	1·6245
	5	27·43	4·29	3·19913	5·20051	1·6256
	6	36·07	6·29	3·19726	5·20258	1·6272
	7	45·73	9·13	3·19456	5·20521	1·6294
MgSn	1	0·87	0·18	3·20247	5·19921	1·6235
	2	1·93	0·40	3·20222	5·20040	1·6240
	3	4·33	0·92	3·20195	5·20189	1·6246
	4	9·05	2·00	3·20114	5·20441	1·6258
	5	11·00	2·47	3·20072	5·20565	1·6264
MgPb	1	0·84	0·10	3·20273	5·19995	1·6236
	2	2·50	0·30	3·20288	5·19987	1·6235
	3	4·03	0·49	3·20309	5·20181	1·6240
	4	11·36	1·48	3·20420	5·20618	1·6248
	5	20·07	2·86	3·20575	5·21447	1·6266
	6	23·33	3·45	3·20620	5·21808	1·6275
	7	26·39	4·04	3·20658	5·22031	1·6280
	8	28·30	4·43	3·20676	5·22285	1·6287
	9	30·81	4·97	3·20705	5·22557	1·6294
	10	38·06	6·73	3·20786	5·23298	1·6313

Further studies on active nitrogen`

III.* Experiments to show that traces of oxygen ~~or other~~ impurity affect primarily the walls of the vessel, and not the phenomena in the gas space

By Lord Rayleigh, F.R.S.

(*Received* 3 *September* 1941)

This paper deals in the first place with the effect of minute gaseous impurities which have apparently an important effect in promoting the formation of active nitrogen by the electric discharge. Oxygen is one of them.

It is shown that the action of oxygen must be on the walls of the tube: for when a minute oxygen tributary is added to the nitrogen stream, it takes much longer to assert its action than the time required to establish the changed composition of the gas stream: and similarly when the oxygen tributary is checked its activity persists for a time. This is interpreted to mean that the action of the oxygen or other impurity is to modify the glass wall in a way favourable to the accumulation of active nitrogen.

The restoration of the afterglow by addition of tributary stream of oxygen can be observed even better with the electrodeless discharge at low pressure (0·3 mm. say) than with high-pressure electrical discharges at say 3 mm. pressure. The intensity in some experiments has been raised 32-fold by the admission of oxygen. Under certain conditions of the glass the cutting off of the oxygen tributary causes a temporary fall of intensity to considerably less than the static value, followed by a rise to that static value. This shows that the phenomena are complicated and not likely to admit of any very simple analysis.

The paper goes on to examine how various treatments of a glass vessel affect its action on the afterglow. As found by Herzberg strong preliminary heating in vacuo makes it destructive of the glow. It is now shown that heating in nitrogen at atmospheric pressure does the same. Heating in oxygen, even at 1 mm. pressure, restores it. It seems clear that these effects cannot easily be explained by the formation or removal of gas layers.

The behaviour of glass is therefore very complex, and not likely to be easily unravelled. The behaviour of the gas itself away from any surface is more fundamental. By a special device it is shown that in this case the purest nitrogen is the best, and that the addition of a trace of oxygen has no favourable effect in promoting the active nitrogen phenomena.

1. Introduction

Many investigations have been made directed to the question of whether the phenomena of active nitrogen can be obtained in pure nitrogen, or whether an admixture is necessary: and if an admixture is necessary, what is its function. In spite of the large amount of care and labour which has been devoted to this matter, I have not been able to feel that the essential questions have been answered. A full summary of the literature will not be attempted, but I will mention the following points, which appear crucial.

---

* I and II. *Proc. Roy. Soc.* A, **176**, 1–27. **August 1940.**

In my own early work, I felt convinced that nothing but nitrogen was involved. The main reasons for this were: (1) That while the commercial nitrogen of that time did not give the afterglow well, it became good by prolonged standing over moist phosphorus to remove the residual oxygen, the only likely impurity. It was, of course, dried after this treatment. (2) The discovery by Fowler and myself (1911) that the characteristic visual afterglow bands belong to the first positive group of nitrogen. No one before or since has ever doubted that these first positive bands are due to nitrogen and nothing but nitrogen. It would nowadays make complete havoc of the successful theory of band spectra to attribute them to anything else. (3) I found that the glowing gas reacted with cold mercury to form nitrides, and with hydrocarbons to form hydrocyanic acid, chemically detected (1911). I could see no place for anything but nitrogen in this. However, the next point appeared to put a different aspect on the matter. (4) It was found by Comte (1913) and by Tiede and Domcke (1913) that the glow could, under certain conditions, be reduced to little or nothing, when oxygen was perfectly removed, and that it could be restored by the intentional admission of oxygen. (5) I found (Strutt 1915) the restoration of the glow could be achieved by the admission of a trace of gases other than oxygen, such as methane or hydrogen sulphide. This showed that it was incorrect to regard the phenomena as necessarily dependent on some reaction between oxygen and nitrogen. I regarded them as conditioned in some way by a catalytic action of one of the impurities introduced, and thought that this probably influenced the discharge mechanism which made the gas active.

2. GENERAL

All these latter experiments under (4) and (5) are very difficult to repeat with certainty, though they have succeeded on many different occasions, with different sets up of apparatus and different methods of purifying the nitrogen used, and in the hands of several different experimenters. In taking up the subject again many attempts were made to find some source of nitrogen which, when passed through a low-pressure discharge apparatus, would not give the glow under any circumstances. These attempts were not successful.

Real progress began when I settled down to the use of nitrogen from one source, and studied quantitatively the luminosity of the glow under varied conditions. This nitrogen was supplied in cylinders by the Osram Lamp Company of Hammersmith. It had, I understand, been purified by passing over hot copper at a high pressure, and was used straight from the cylinder, after passing at a pressure slightly above atmosphere through a column of phosphorus pentoxide 2 cm. diameter and 60 cm. long.

Tests of purity. This nitrogen did not show any trace of luminosity when bubbled through a solution of phosphorus in pure olive oil. This goes to show that there is no appreciable amount of oxygen in it. Further, when the afterglow spectrum from it is photographed through a quartz window, no trace of the β and γ bands (Strutt 1917) appears. These bands are a fairly sensitive test for oxygen or carbon dioxide, and come in very definitely when oxygen is admitted to the extent of 0·1 %.

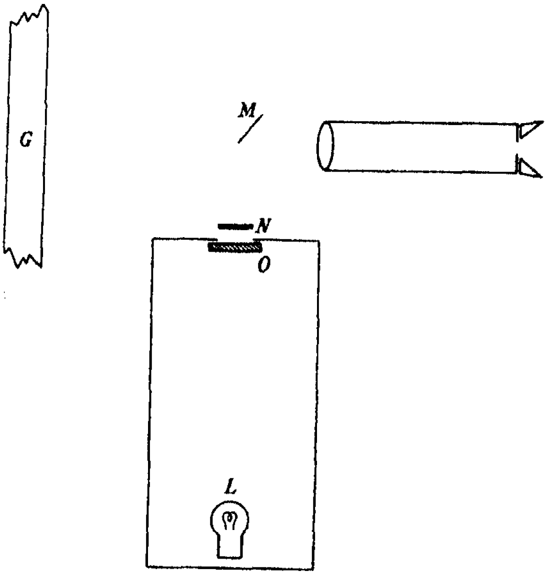


FIGURE 1

Further, there was no trace of the cyanogen spectrum in the afterglow which is a sensitive test for hydrocarbons or for carbon monoxide (Strutt 1911). The afterglow spectrum in fact shows nothing but the α bands and the resonance line 2537 of mercury, derived from the mercury used in the pressure gauge and in the mercury-vapour pump. Several hours exposure were given, working at 2·3 mm. discharge pressure.

Photometric arrangements. The photometry in this work was all relative to an arbitrary standard. Since it was designed for studying the effect of various changes in the conditions, there was no need for anything more.

The arrangement is shown diagrammatically in figure 1. The afterglow light *G*, in a tube of 2 cm. diameter, is viewed directly though not in focus. It is compared with a light from a small incandescent lamp with controlled voltage *L*, diffused by opal glass at *O*. The photometric edge is afforded by a silver-on-glass mirror at 45° M, which is viewed in focus by a lens of 5·7 cm. focal length, through an ocular diaphragm at 10 cm. distance from the lens. The lens and diaphragm are carried in a brass tube. The lens acts to some extent as a field lens, and its area affords the field of view, bisected by the edge of the silver-on-glass mirror, which is seen as a horizontal diameter.

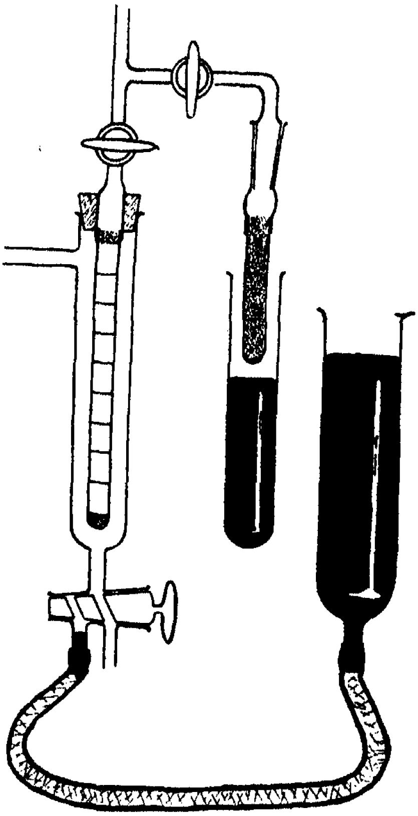


FIGURE 2

The light under investigation is equalized to the standard by one of a series of neutral glasses indicated at N. These glasses are arranged along radii of a rotating disk, with vertical axis (not shown) which brings them successively into position. A well-marked click indicates by sound and touch when each glass falls into position, and a legible index in front shows its number. These numbers are 0–13, and denote 10 times the density, i.e. $10 \times \log_{10}$ of the opacity. Thus, for example, glass no. 7 has a density of 0·7 and an opacity of 5·01. The range of thirteen glasses is not always enough. If the source is too bright, a supplementary glass of known number is placed over it, and the resulting number is negative.* If glass 13 still leaves the comparison light, too bright a supplementary glass is placed over the latter and its value is added. The disk arrangement was built by my son, Mr G.R. Strutt, largely of Meccano parts. The advantage of it is that the value can be rapidly varied in the dark, clicks being counted to determine the change made.

---

* As in the case of magnitude number of the brightest stars.

The nitrogen stream was controlled, as before, by a pipeclay tube (figure 2), sealed at one end, through which the gas passed from the high pressure (15 cm. above atmosphere) into the vacuum system. The pipeclay tube was graduated in centimetres, and a greater or less length could be used by partly immersing it in mercury. The rates of intake were determined by means of a graduated vessel. If the mercury is lowered, the three-way tap at the bottom allows gas to be blown off to wash out the space round the pipeclay tube.

To admit a trace of air at pleasure to the nitrogen stream, a second porous pipeclay valve was used, the area of porous surface exposed being much smaller. The pipe stem was sealed at one end as before, and it was found convenient to cover it all over with sealing wax in the first instance. The sealing wax was removed locally by means of a small drill held in the fingers. In this way five holes were made in a vertical row, each somewhat less than 1 mm. diameter and 1·5 cm. apart. These could be opened in succession by lowering the level of mercury in which the tube was immersed. There were thus five stages. The rate of admission at each of these stages was readily determined once for all from the rate of use of pressure on a known volume. To start the air tributary, the mercury test-tube covering it was lowered to a fixed stop. To shut it off, the test-tube could be raised again. But in some cases a stopcock placed as close as possible to the main stream was used, to ensure a sudden cut off. The results were, however, sensibly the same whichever method was adopted.

3. OXYGEN PRODUCES AN EFFECT ON THE WALLS

If, as the earlier workers tacitly assumed, the glow phenomena were directly conditioned by the presence of a small volume concentration of oxygen in the nitrogen, then it is evident that the glow intensity must increase immediately this condition is established, and it must disappear again immediately the condition is removed. I have long suspected that the effect showed itself much more slowly but have lately brought the question to as sharp a test as possible. The plan of experiment was as follows: Pure nitrogen was passed at a definite rate and at a definite pressure through a discharge tube followed by an observation tube. The intensity of the glow was measured in the steady state. A definite small percentage of air was admitted at time zero, and the consequent variations of intensity with time were determined. In this way the time to establish a substantial increase of intensity was found. It could be compared with the time required to alter the composition of the gas by sweeping it out of the tube system at the known rate of streaming. A similar test could be made from the time when the air admission was cut off.

For this purpose it was necessary to design the apparatus for constancy of operation. The gas intake was regulated by the pipeclay diffusion valve, and a mechanical pump working at constant speed kept the gas pressure constant. The discharge was by a 50-cycle a.c. This is much less effective than the condensed discharge for producing intense afterglow, but high intensity was not necessary or even desirable for the work in hand. The advantage of the low-frequency alternating current is that it lends itself to steady and uniform working, and this can be checked by ordinary measurements of r.m.s. current and terminal voltage.* The discharge tube was constructed as in figure 3. It will be noticed that the gas stream is divided into two equal parts, each of which flows through a tubular electrode of sheet nickel. The two halves of the stream reunite and pass through a cranked connexion into the observation tube when the glow intensity can be observed photometrically. The cranked connexion is to stop false light from the discharge. It is blackened externally.

The reasons for this peculiar construction are as follows: It is essential that the gas should pass straight through the tube, so as to flush it out in a determinate time when the gas composition is altered. This condition will not allow the use of side tubes for the electrodes: for such side tubes would act as 'backwaters' in which the stream would be stagnant. It is necessary therefore to place the electrodes in the main stream. The more obvious construction is to have a straight discharge tube with the electrodes placed co-axially in it as in figure 4. This would probably succeed if a condensed discharge were used, but with a heavy uncondensed discharge no afterglow is obtained from such a tube. The electrodes get hot, and the afterglow is destroyed by contact with the electrode near the exit. It was necessary therefore to draw the gas off from a mid-point between the electrodes. The straight tubes at the top (figure 3) of 1 mm. bore divide the gas equally between the two branches, and are so narrow that the discharge has no inclination to pass through them. An electric fan was used to keep the discharge tube cool.

An experiment with this arrangement will now be quoted in detail.

Nitrogen pressure in discharge tube 2·2 mm. Nitrogen admitted per second 0·3 c.c. at atmospheric pressure, equivalent to 104 c.c. at low pressure.

Volume of discharge tube and connexions from air inlet to photometer 95·5 c.c.

---

* High-voltage continuous current is of course even better, but the facilities for it were not available.

(It may be added that the discharge current was 33 mA and that the air tributary was $5{\cdot}90 \times 10^{-3}$ c.c./sec., equivalent to 0·4 of 1 % of oxygen introduced into the nitrogen stream. These data do not enter into the calculation however.)

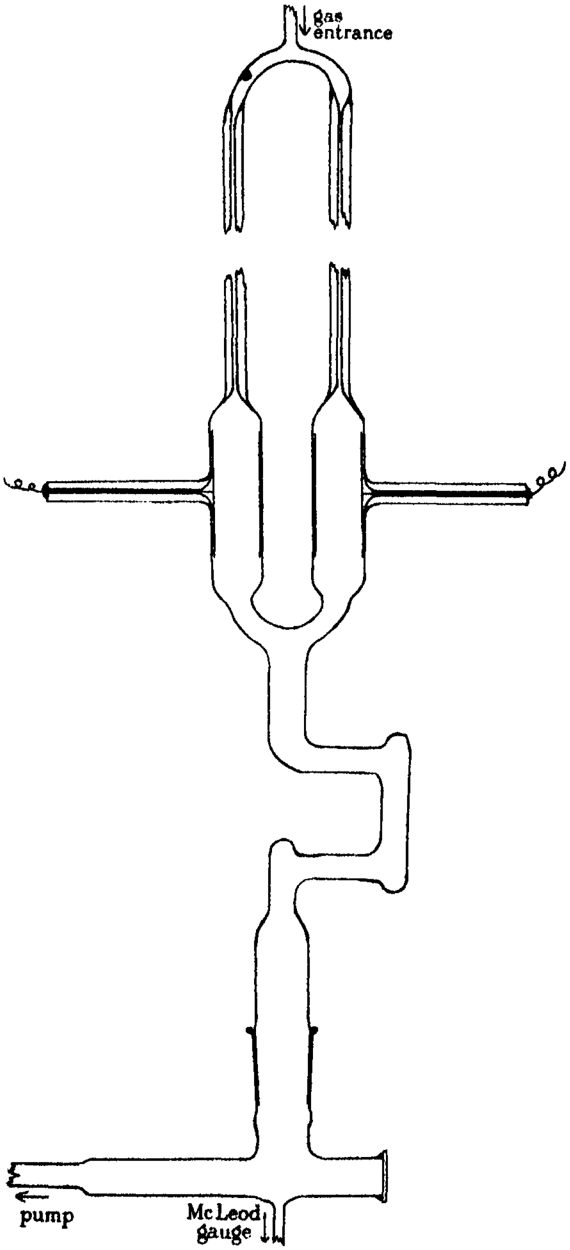


FIGURE 3 (¼ actual size)

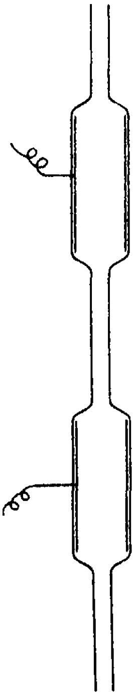


FIGURE 4

Time required to wash out the discharge tube 95·5/104 sec. or 0·91 sec.

[This time was approximately verified by the direct method of observing the disappearance of the afterglow. When the discharge current was switched off the glow was no longer generated, and its fairly abrupt after-boundary was seen travelling along the tube. It passed the place where the photometer was applied in a time judged to be less than 1 sec. after switching off.]

The following table gives the result of three runs, starting from the steady state attained when the air tributary is left on. The air tributary is suddenly cut off at time 0, and the intensity decreases. The runs were then made in the reverse manner, starting from the steady state without an air tributary, and suddenly putting it *on* at time 0.

photometer reading	intensity	time sec.	mean
0	7·9	0, 0, 0	0·0
2	5·0	3, 4, 4	3·7
4	3·2	10, 11, 12	11·0
6	2·0	18, 18, 24	20·0
7	1·6	29, 33, 40	34·0
8	1·3	55, 44, 49	49·3
9	1 (assumed)	100, 90, 74	91·0

photometer reading	intensity	time sec	mean
9	1 (assumed)	0, 0, 0	0·0
7	1·6	8, 8, 8	8·0
5	2·5	11, 10, 10	10·3
3	4·0	13, 12, 12	12·3
2	5·0	20, 14, 16	16·7
1	6·3	28, 22, 24	24·7
0	7·9	65, 58, 61	61

It is clear that these times are very long in comparison with the time 0·9 sec., required to flush out the tube, and to remove, or establish, the small oxygen concentration added by the tributary stream.

Naturally, when the limiting brightness is being asymptotically approached, the time of reaching it is not well defined. This accounts for the considerable discrepancies at the end of the successive runs, and makes it undesirable to state, or attempt to compare, the respective total times taken for decay or for recovery. It is better to take a part of the range only.

If we take the intensities 5·0 and 1·6, the time taken to pass over this interval is 30·3 sec. when the intensity is falling, and 8·7 sec. when it is rising. Thus getting rid of the glow by cutting off the air definitely takes longer than recovering it by adding air. [The process of cutting off the air tributary by a tap is more satisfactorily abrupt than admitting it by uncovering a porous tube: nevertheless, the former takes longer to assert itself, so that the uncertainty, if any, cannot affect the conclusion that the recovery process on adding air is the quicker.]

That the effect of added oxygen persists for a relatively long time after the added oxygen has been removed from the gas space seems to allow of only one interpretation—its effect must be to modify favourably the walls of the tube. To be more specific, it seems natural to assume that the admission of a trace of oxygen, of the order of a fraction of one per cent of the nitrogen, leads to the building up of a layer on the glass which has a favourable effect.

If we are right in assuming that such a layer exists, it seems to be formed, and also to pass off when the oxygen tributary is checked, equally well whether the discharge is going or not. The following experiments may be cited in evidence of this.

The current was maintained throughout at 43 mA.

Pressure 2·2 mm.

Nitrogen stream 0·272 c.c./sec.

Oxygen tributary $1{\cdot}34 \times 10^{-3}$ c.c./sec.

After prolonged running:

	photometer reading	intensity
oxygen off	13	1 (assumed)
oxygen on	1	16

These are the values in the steady state, with the current running continuously. It was found that, having the oxygen tributary on for several minutes but the electric current off, except for the 2 sec. or so necessary for each test, the reading remained at the value 1, intensity 16.

It appears therefore that the oxygen tributary maintained the glass in a condition to give the glow, whether the electric current is running or not.

The rate at which the effect passes off when the tributary is stopped is also independent of whether the electric current is passing. In the following tests the tributary was cut off by a tap at time 0:

	photometer reading	intensity	time min.	time sec.
electric current running	0	20	0	0
	1	16	0	20
	3	10	0	45
	4	7·9	1	45
	5	6·3	4	30
	6	5·0	7	0
electric current not running	1	16	0	0
	5	6·3	4	30
electric current not running	1	16	0	0
	5	6·3	4	30

It may be noticed in passing that in this series of tests the effect of the oxygen tributary passes off much more slowly than in the previous set. The discharge tube was not the same—it had been replaced after an accident—but its construction was similar. The exact quantitative reproduction of these phenomena has, I must confess, proved to be beyond my control, though they admit of fairly exact quantitative repetition again and again, with a given set up of apparatus.

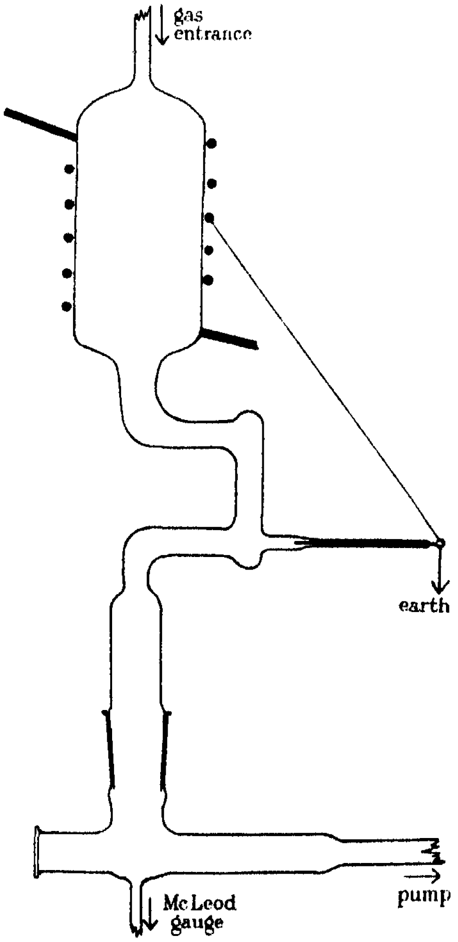


FIGURE 5 (⅓ actual size)

The experiments made hitherto on the restoration of the nitrogen afterglow by addition of a trace of oxygen have all been made at relatively high pressure of 2–4 mm., using a discharge with electrodes. It was considered desirable to try the same thing with the electrodeless discharge which is specially effective for producing the afterglow at low pressure of 0·1–0·4 mm.

The discharge vessel used was a cylinder (figure 5) 5 cm. diameter and 9 cm. long. As in other cases, a cranked connecting tube, blackened externally, was used to exclude false light from the discharge getting access to the observation space. With this kind of excitation there is another source of false light to be avoided, namely, stray discharge. The trouble could be dealt with by providing a wire electrode as shown, which was connected to earth, and also to the middle of the exciting coil. A slight parasitic discharge was taken by this wire, but none penetrated into the observation space.

The 'Hyvac' mechanical pump used in the previous work was inadequate for maintaining a gas stream at the lower pressure, and it was supplemented by a two-stage Gaede mercury-vapour pump.

It was found difficult at first to get the glow down to a low intensity, using the Osram cylinder nitrogen as before. It was ultimately done by prolonged streaming with the discharge going, best if the fan normally used to keep the discharge tube cool was stopped. This, and the work of Herzberg (1928), suggested that heating alone might be effective. The discharge tube was therefore made as hot as was judged prudent by a resistance heater, the exciting coil being removed for the time. The baking out was continued for some time, and the discharge started. It was then found that a marked increase of intensity occurred when the tributary stream of air was admitted, and that the original low intensity was ultimately restored when the air was stopped.

Thus, using a nitrogen intake of 0·098 c.c./sec. and an air tributary of $6{\cdot}69 \times 10^{-3}$ c.c./sec., the oxygen introduced was 1·4 % of the nitrogen. The discharge pressure was 0·32 mm. After baking out for 6 hr. at a temperature estimated at 400° C, the results were:

	photometer reading	intensity
initial	14	1 (assumed)
air admitted	−1	32
final	14	1

The above was a typical result when the discharge vessel had been well blacked out, and the initial intensity had been got to a very low value. When, however, the initial intensity was not so low, the effect of starting the air tributary was to raise the intensity as before, but on cutting off the air tributary a very peculiar effect was noticed. The intensity fell, and it fell much below the steady value at which it had stood initially, slowly rising again to this value.

This was a striking effect, which could be repeated again and again as often as desired, if the tube was in the right condition. How to get the tube into this condition is a matter which has not yet been sufficiently clarified by repeated experiments. The graph (figure 6) shows a typical experiment of this kind. In this case the nitrogen intake was 0·118 c.c./sec. and the air tributary $2{\cdot}65 \times 10^{-3}$ c.c./sec. Thus the oxygen introduced is 0·45 % of the nitrogen. The discharge pressure was 0·4 mm.

Calling the initial intensity unity, it rises to 6·3 on admission of air. On cutting the air off it falls to 0·20, and then in the course of about 3 min. it rises again to the initial value of unity. No explanation of this behaviour can be offered at present. It affords further evidence that surface effects are involved, and that these are highly complex.

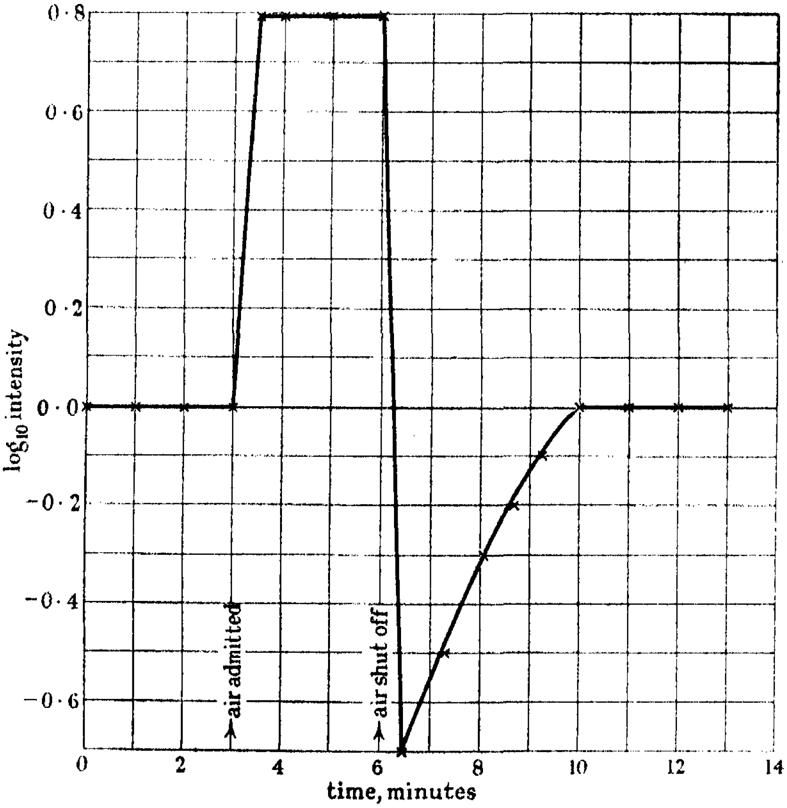


FIGURE 6

The fact that the oxygen tributary acts in just the same way whether the electric current is passing or not seems to show clearly that the formation of oxides of nitrogen by the discharge, which might be adsorbed on the glass surface, has nothing to do with the matter.

As to what the effect of oxygen is and why it acts in the manner described I have no definite suggestion to offer.

4. EFFECT OF GLASS SURFACES VARIOUSLY TREATED ON ACTIVE NITROGEN

It may be felt that more evidence is required before we can admit that a clean glass surface can be in such a state that it is strongly destructive of the glow. Since well-cleaned glass vessels are in general the most favourable, and since a newly made glass apparatus when first evacuated takes time to get into a favourable condition and improves with prolonged drying, it seems rather paradoxical to assume that clean glass can be destructive.

Nevertheless, the investigations of Herzberg (1928) decidedly suggest that this may be the case. He found that a silica discharge vessel which had been strongly heated got into a state when the glow could not be obtained in it even after cooling, using the electrodeless discharge and static condition (no gas stream). I have made a modification of Herzberg's experiment which exhibits the effect in a specially convincing way, since treated glass can be compared with untreated at a single *coup d'œil*, everything being strictly symmetrical between them.

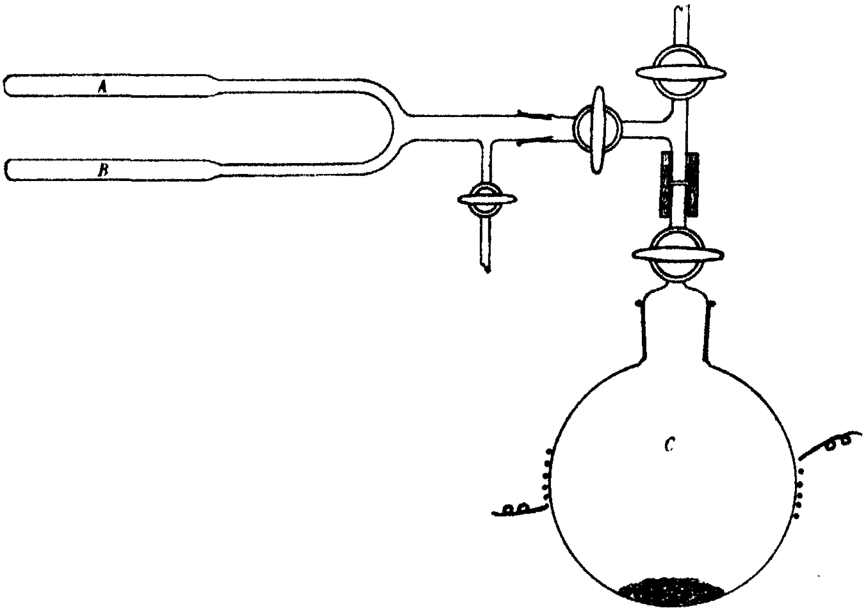


FIGURE 7

A tube system *AB* was constructed of pyrex glass as shown in figure 7. It was exhausted and the upper branch *A* was heated by a resistance furnace as strongly as possible, short of sucking in. The temperature was 575° C. This was continued for an hour or more and *A* was allowed to cool. On admitting glowing nitrogen from the stopcock from a large electrodeless discharge bulb *C* at 0·3 mm. pressure, the glow was bright in the lower branch *B*, but scarcely if at all visible in the upper one *A*, being evidently destroyed by the walls as fast as it could enter.

By keeping up the supply of active nitrogen the glow in *A* gradually became perceptible, and after prolonged action it ultimately became as bright as in *B*. By heating *B* it in turn could be made dark, and a corresponding cycle of changes could be observed.

Herzberg (1928, p. 888) found that the introduction of a trace of hydrogen into electric discharge was the condition for getting the glow back in his silica bulb. In the present work the glow could be got back again without this. Introduction of hydrogen through a palladium osmo-regulator was tried, and may have accelerated the action somewhat, though I cannot state positively that this was the case.

The tube *A* was baked out again, allowed to cool, and left exhausted for a week. On admitting the glowing gas at the end of that period, *A* was dark while *B* was bright. There is thus no spontaneous recovery, or at all events none appreciable in this period of rest. Nor does having the tubes full of air or carbon dioxide at atmospheric pressure and temperature for 24 hr. produce any recovery.

On filling with ammonia gas for 24 hr., the recovery if any was slight. This experiment was suggested by Herzberg's view that a trace of ammonia on the glass surface is the favourable condition for glow. He supposes that the ammonia is removed by the baking process. The experiment, however, is not favourable to this explanation.

The tubes were next filled with distilled water, and left overnight. On pouring it off, exhausting in presence of phosphorus pentoxide, and testing, the baked tube was found to be considerably restored, though even then it was by no means equal to the other one.

In the experiments with this apparatus so far described the glowless condition was obtained by heating the pyrex tube in vacuo. I find, however, that the tube can also be rendered glowless by heating it for an hour or two in hydrogen or nitrogen at atmospheric pressure. The nitrogen used was freed from oxygen by standing moist phosphorus in the cold, and drying with phosphoric oxide. On the other hand, heating in oxygen or in air will not make it glowless, and if it is glowless initially it gives the glow again in full brightness after this treatment.

This effect is shown even by oxygen at 1 mm. pressure, with a few hours heating.

Heating in carbon dioxide at 760 mm. also restores the glow, and after this treatment it is very difficult to get the tube glowless again.

I have repeated all these effects with a tube of silica.

The above statements are soon made, but much time has been spent in repeatedly verifying them. There are one or two traps to be guarded against. One of these is that pyrex tubes are reduced by hydrogen, with formation of a brown deposit, probably arsenic (Randal and Leeds 1929). After this becomes apparent I have not been able to put the tube into the glowless condition again, but a new tube shows the normal effect. Again, I have found that a silica tube, when taken into use, sometimes shows anomalous effects. After softening it in the blowpipe, its behaviour becomes normal and it will then show the above-described effects with heating in vacuo and with various gases, in any order, and always with the same results, except for the difficulty of getting it dark after it has once been heated in carbon dioxide.

It is difficult to see the full bearing of these results. They do not favour the view that freedom from occluded gases will explain the glowless condition, for having the tube full of air has no effect. Heating in hydrogen produces the glowless condition, and heating in oxygen restores the glow. If this were all we might invoke chemical reduction as the cause, reversed by oxidation. But then the behaviour of nitrogen in inducing the glowless condition, and of carbon dioxide in restoring the glow, would remain unclassified.

The experiments described in this section are, it is thought, of considerable interest in themselves. It is not clear, however, that they bear very directly on the tests when the glow is restored by a little oxygen added to pure nitrogen; for in the present case the gas is not initially pure, and treatment with cold gases even at atmospheric pressure does not restore the baked tube to its normal condition. Further work is required to elucidate the relation between these phenomena and the others. The broad general fact which emerges is that an apparently clean glass surface may be brought into a condition which is destructive to the glow, and after all this is the main point. However curious the behaviour of glass may be, the more fundamental question concerns active nitrogen itself rather than its relations to the containing vessel.

5. EXPERIMENTS IN WHICH THE WALL EFFECT IS ELIMINATED

The experiments so far described show clearly enough that the observed restoration of the afterglow at low pressure when a trace of oxygen is added to the nitrogen in use is an effect on the walls of the tube; for it does not assert itself until some time after the new composition of the gas is established. I wished, however, if possible to examine the effect of adding oxygen under circumstances when the action of the walls of the vessel would be eliminated entirely.

At the low pressure used to get the afterglow in its most conspicuous form this is impracticable, because diffusion is so rapid at these low pressures that the glowing gas gets to the walls immediately. At higher pressire, however, diffusion is not so rapid, and phenomena can be observed which are clearly unconnected with the walls.

After preliminary trials, the arrangement of figure 8 was adopted. A pair of bare nickel wires are arranged radially in a bulb of 500 c.c. capacity, and converge to within a distance of about 7 mm. apart. These wires are bare without any kind of sheath. The nitrogen pressure in the bulb is maintained at 6 cm., and the condensed discharge from a small coil with jar and spark cap passes between these wires where they appear most closely at the centre of the bulb. Under these circumstances the spark is surrounded by an aureole of orange colour, showing the characteristic afterglow spectrum and remaining luminous for 1 sec. or so after the discharge is turned off. Thus this aureole visibly takes its origin from the spark, and owing to the local character of the excitation it is very much weakened by dilution before it reaches the walls of the vessel. Moreover, it comes off from the spark in visible streamers, and by no means with uniform spherical distribution. These streamers show the turbulent motion of the gas, and their movements are capricious. The point is that their progress from the spark towards the wall of the tube shows divergence, so that there is no kind of doubt that what is observed near the spark is independent of the wall of the tube.

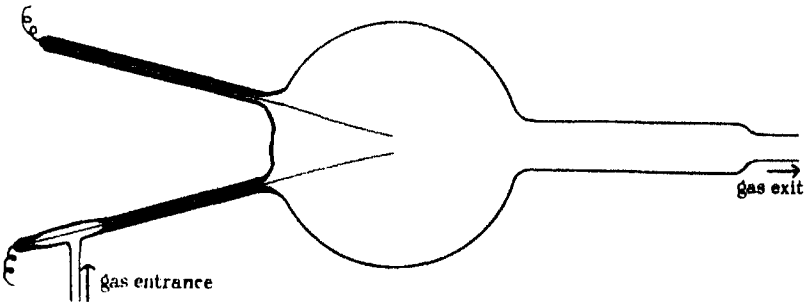


FIGURE 8 (⅓ actual size)

These phenomena do not absolutely require a stream of gas passing through the bulb, since we depend on diffusion and the convection currents set up by the heat of the spark to separate the glowing gas from the discharge. However, it was thought best to use the same technique as before for introducing a trace of air into the nitrogen, and this method involves the use of a stream of pure nitrogen into which a small tributary stream of air is admitted. Osram cylinder nitrogen was used as before, dried by phosphorus pentoxide. The rate of intake was similar to that in the low-pressure experiments, and to maintain 6 cm. of pressure the action of the pump was suitably choked by a capillary tube. With these arrangements the streamers were kept under observation during a long run, so as to be sure that all impurities were washed out and the standard conditions established. The development of the streamers reached a definite limit, and did not change for hours. They reached as far as the glass wall in this case, though, as already explained, their origin at the spark was obvious. The nitrogen intake was 916 c.c./hr., measured at atmospheric pressure.

An air tributary was then fed in, amounting to 7·56 c.c./hr., thus introducing oxygen to the extent of $1{\cdot}65 \times 10^{-3}$ of the nitrogen. The effect of this was extremely marked, and as long as it continued the extent of the streamers was much reduced, approximately as shown in figures 9 and 10. This was tried on and off repeatedly. It took some minutes for the effect to assert itself fully, because of the slowness of the gas stream, and some minutes to pass off completely when the air admission was stopped, but the result was perfectly definite, and proved to my complete satisfaction that, at a pressure of 6 cm. and when the action of the walls is eliminated, the admission of oxygen to this small extent, so far from developing the afterglow streamers, conspicuously diminishes them.

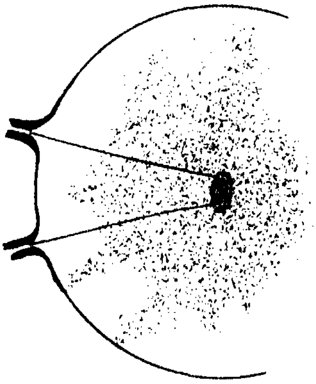


FIGURE 9

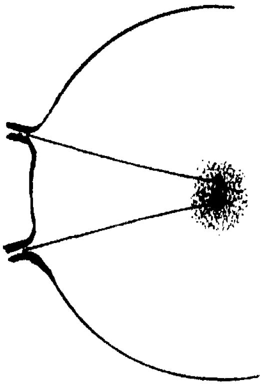


FIGURE 10

Further tests were made with an air admission of 3·21 c.c./hr., giving oxygen admixture of 7×10^{-4} of the nitrogen. The effect of this was still markedly unfavourable: but the figures are drawn in the attempt to indicate the more conspicuous effect of the larger admixture. Unfortunately, the experiment does not lend itself to a quantitative statement of the effect.

REFERENCES

Comte, F. 1913 *Phys. Z.* **14**, 74–76.
Fowler, A. and Strutt, R. J. 1911 *Proc. Roy. Soc.* A, **85**, 377–388.
Herzberg, G. 1928 *Z. Phys.* **46**, 878–895.
Lewis, B. 1929 *J. Amer. Chem. Soc.* **51**, 654–674.
Randall, J. T. and Leeds, R. E. 1929 *Trans. Soc. Glass Techn.* **13**, 16–19.
Strutt, R. J. (Rayleigh) 1911 *Proc. Roy. Soc.* A, **85**, 219–229.
Strutt, R. J. (Rayleigh) 1915 *Proc. Roy. Soc.* A, **91**, 303–318.
Strutt, R. J. (Rayleigh) 1917 *Proc. Roy. Soc.* A, **93**, 254–267.
Strutt, R. J. and Fowler, A. 1912 *Proc. Roy. Soc.* A, **86**, 105–117.
Tiede, E. and Domcke, E. 1913 *Ber. dtsch. Chem. Ges.* **46**, 4065.

Further studies on active nitrogen

IV. The ionisation associated with active nitrogen

BY LORD RAYLEIGH, F.R.S.

(*Received* 3 *September* 1941)

In this paper the ionization associated with afterglowing nitrogen is examined more thoroughly than has yet been done. It is found that this ionization is completely cut off if the testing vessel is separated from the afterglow by a silica wall. The ionization cannot therefore be directly produced by light, at any rate within the range from λ 1850 to the visible. No increased current is observed when the surface action of the active gas is so vigorous as to make the testing cathode (of gold) red hot.

The ionization is in general nearly the same when the cathode of the testing vessel is one of the common metals or a surface coated with metaphosphoric acid. Copper, however, is an exception. While the copper is clean it gives an effect several times larger, though this soon goes off, as it becomes dull by the action of the gas. It appears therefore that though there is some surface emission of electrons, it is in most cases small compared with the volume ionization in the gas space.

Comparisons are made of the number of ion pairs generated, and the number of photons emitted, per cubic centimetre at various stages as the glow dies down. At first the number of photons is many times greater, but as the glow gets down the numbers become equal, and at intensities too low for convenient observation it is likely that the number of ion pairs may become greater than the number of photons. It is certain that the light emission and the ionization do not go down at the same rate. The latter goes down much more slowly, and gains relatively. The interpretation of this is discussed.

It is found that the admission of inert nitrogen which is known to increase the (instantaneous) emission of photons, also increases the (instantaneous) ionization.

It is pointed out that the ionization of nitrogen requires 15·51 V, and the presence of this ionization seems to throw much doubt on the attempts to calculate the energy of active nitrogen from the band spectrum emitted, which suggests an energy of 9·6 V.

1. INTRODUCTION

In early papers (Strutt 1911, 1912) it was shown that a strong ionization current can be observed between electrodes immersed in the nitrogen afterglow. The current collected was found to be very much smaller than would correspond to the electrochemical equivalent of the active material, as determined by chemical methods. It was shown by a variety of considerations that this current is not due to the survival of ions from the generating discharge. For example, other gases do not give any appreciable ionization current under similar conditions. It is only obtained in presence of the peculiar nitrogen afterglow. Other phenomena to be presently discussed (see § 6 below) are equally conclusive on this point.

In this work I had taken it for granted that since the luminous phenomena obviously originated in the volume of the gas, the ionization also originated there. Constantinides (1927) found that using clean brass electrodes, arranged as concentric cylinders, the saturation current was greatest when the outer electrode was negative, and that the current was proportional to the area of the negative electrode. He concluded from this that the effect originated at the negative electrode. He left it open whether this was simply due to photoelectric action on the electrode by the light given off by the glow, or whether it was due to the liberation of electrons by the direct action of impinging atoms.

2. TEST FOR A PHOTOELECTRIC EFFECT

To test as far as possible whether there was any direct photoelectric effect on the electrodes, or rather the negative electrode, the following experiment was tried.

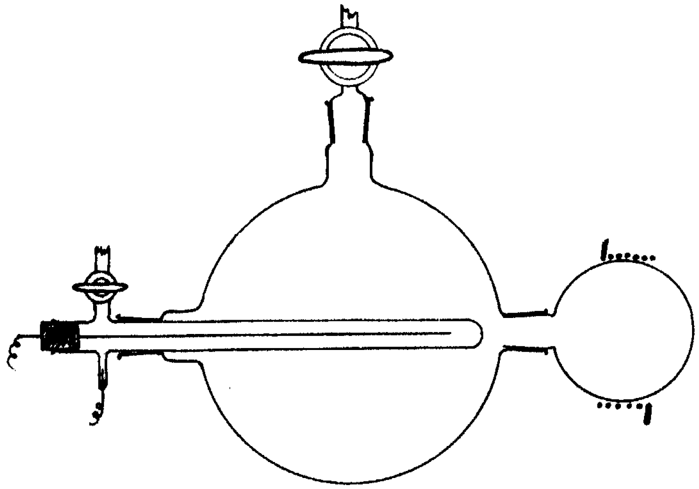


FIGURE 1 ($\frac{1}{5}$ actual size)

A 5 l. globe as shown (figure 1) was arranged with a silica tube about 1 mm. thick, and 18 mm. inside diameter, projecting right into it as shown, to a distance of 15 cm. The silica tube was smeared internally with metaphosphoric acid to make it conducting, and contained a central electrode of brass wire, insulated from the coating. The central electrode and the phosphoric coating were connected up in series with battery and electrometer, the latter used as a galvanometer with a high-resistance leak to earth, so that 1 scale division deflexion represented $1\cdot44 \times 10^{-11}$ amp. The gas pressure both inside and outside the silica tube was 0·2 mm., but these spaces were not in communication, except by the light, visual or ultraviolet, getting through the walls of the silica tube. It was proved in an independent test that the walls as coated with metaphosphoric acid were in fact transparent as far as the aluminium line λ 1850. If the action had been due to light within this limit, it would have had almost as good a chance of acting as if the silica tube had been away.

When the globe was filled with very bright nitrogen afterglow, not the slightest deflection of the electrometer could be detected. If the silica tube had been absent, and the glowing gas had had direct access to the brass electrode, an enormous effect would have been produced, certainly not less than 10^4 times any that can in fact have been present, and probably very much more still. I conclude, therefore, that direct photoelectric action by any wave-length greater than λ 1850 is excluded. There is still of course the possibility that photoelectric effects by smaller wave-lengths than this might be operative. Such wave-lengths would not get through the silica tube. There is no evidence however that active nitrogen emits these shorter wave-lengths.

3. A GOLD CATHODE GIVES NO INCREASED IONIZATION WHEN CAUSED TO GLOW IN ACTIVE NITROGEN

It was shown previously (Rayleigh 1940, p. 16) that gold and other metals can become red hot in active nitrogen. When this occurs it may be assumed that the action of the gas on the surface is at its maximum, and if the ionization were in the main due to the emission of electrons in a surface action it would be expected to increase greatly when the gold becomes red hot.

The arrangements used for the experiment to test this point are shown in figure 2. In this and other cases where the afterglow is to be examined in presence of metal surfaces, it is impracticable to use the method, much preferable on other grounds, of stopping the discharge, and making all measurements on a supply of active nitrogen generated before the measurements begin. In the presence of glass surfaces only, either clear or, better, coated with sulphuric or phosphoric acid, and where the activity of the gas does not need to be very high, this method can be used. But when a supply of strongly active gas is required to be kept up, or when the active gas is exposed to other kinds of surface, more particularly metal ones, a rapid loss of activity cannot be avoided, and this method ceases to be easily applicable. It becomes necessary to keep up the supply of active gas, generating it while the experiment is in progress. This is conveniently done in a smaller bulb in permanent connexion with the larger one. The electrodeless discharge is kept going in the smaller bulb, the inner walls of which are of course of clear glass, uncoated, and the glowing gas passes into the larger bulb by diffusion.

A piece of gold foil about 1·5 sq. cm. total area (both sides included) was placed in the centre of a 5 l. bulb (figure 2). It was attached by means of a rather fine platinum wire, the two ends of which were brought to the outside. The inside of the bulb was coated with metaphosphoric acid, and this and the gold wire were joined up in circuit with a 9 V battery and a galvanometer. When the bulb was full of active nitrogen generated in the smaller bulb (300 c.c.) which was joined to it, the gold did not as a rule become red hot immediately, and the current could be measued while it was dark. The gold could then be started glowing by momentarily heating the wire by which it was suspended by a current. Once started, the action continued spontaneously. The ionization current between the gold and the coating of the bulb could then be measured again with the gold glowing.

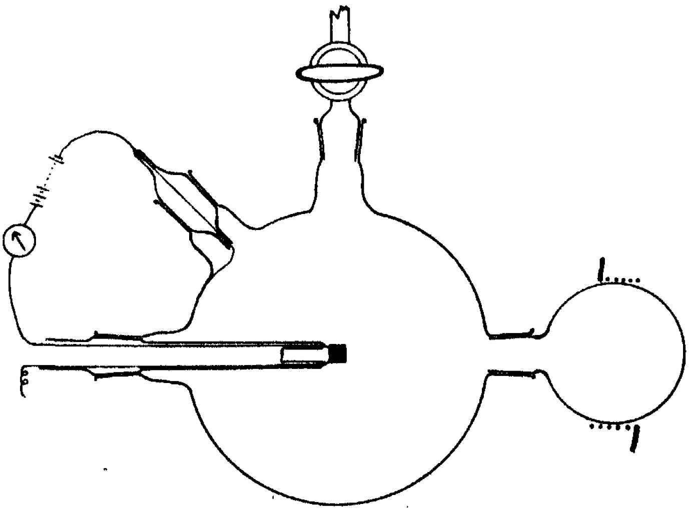


FIGURE 2 ($\frac{1}{5}$ actual size)

After a short rest the gold usually became inactive again, and the observations could be repeated.

The following were the results at various gas pressures:

pressure cm.	current, amp. × 10^{-7} gold dark	current, amp. × 10^{-7} gold glowing
4×10^{-2}	6	5
3×10^{-2}	8	5
2×10^{-2}	12	5
1×10^{-2}	3	2
5×10^{-3}	3	2

It will be noticed that when the gold started glowing the current was always diminished. The experiments therefore do not encourage the idea that the ionization is due to the emission of electrons by the imparting of energy to the metal by the active gas. When energy is being vigorously imparted, as evidenced by the metal getting red hot, the current is not increased, but diminished. It may naturally be asked why this should be the case. It could be anticipated if the ionization is regarded as a volume effect in the active gas; for if the activity is being drawn upon to make the metal red hot, there is naturally less of it left in the gas. This is illustrated by the fact that the afterglow light becomes dimmer. In the last of the experiments cited, at the lowest gas pressure, the afterglow light was almost extinguished when the gold became red hot.

4. Comparison of different electrode metals

The most obvious criterion for deciding whether the action which produces ionization is in the main at the surface of the cathode is to substitute cathodes of different materials, and determine whether or no this affects the current produced. In the case of the photoelectric effect, for example, this criterion at once shows that we have to deal with what happens at the surface, although when it comes to obtaining definite and specific data about different surfaces, the problem has proved far from easy. However, the enormously greater effect of the electropositive metals when they have clean surfaces is easily established.

The arrangement used in the case of active nitrogen is shown in figure 3. The active nitrogen is produced in the small bulb on the right. The large bulb acts as a reservoir. It produces a steadying effect against the irregularities of the discharge, and removes the latter from too near the neighbourhood of the measuring vessel. The vessel is of ebonite, closed at the ends by metal plates, which are made airtight with plasticene applied at the outside.

These plates are in general of two different metals. They are connected up with an 18 V battery and a sensitive galvanometer, with a commutator arranged so that either plate can be made negative. The experiment is then directed to see whether any notable change in the current occurs on commutation. If negative electrons emitted from the cathode play an important part in the ionization, the metals should differ, and a larger current should be obtained from one metal used as cathode than from the other.

As a standard, rustless ('Staybrite') steel was generally used as one of the plates, and another metal compared with it. The metals were carefully cleaned with emery paper or otherwise, before putting them in, and touching with the fingers was avoided. Some of the metals were not available in the form of thick plates suitable to stand the atmospheric pressure. In such cases a second rustless steel plate was used, and a disk of the metal to be tested was allowed to rest on it inside the vessel. When it was desired to test a metaphosphoric acid coating, this coating was applied to a rustless steel plate.

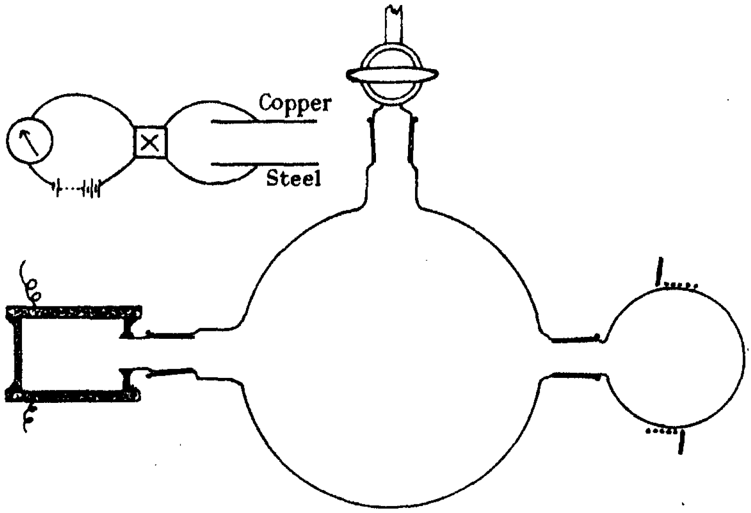


FIGURE 3 (⅕ actual size)

The following table gives the comparative results, the current with 'Staybrite' steel cathode being units:

copper	6·5 down to 1·03	gold	1·11
brass	3·14 to 1·06	tin	1·03
platinum	1·37	zinc	0·99
nickel	1·20	lead	0·98
silver	1·18	metaphosphoric acid	1·00
aluminium	1·14		

The above results are given as representative of a good many experiments that have been made. The results with a given metal were not closely recovered in successive experiments, and the method with spark excitation and its inevitable variation is not capable of giving anything but rather rough results.

It may be said generally that the majority of metals do not give a result notably different from rustless steel used as the standard, and in particular that the coating of metaphosphoric acid behaves like these metals. It would be strange if these varied surfaces all gave off electrons to a closely comparable extent under active nitrogen, and the natural inference is that none of them do so to any important extent, and that the ionization has its origin in the gas space.

The case of copper (and brass which contains copper) is different. This metal when clean is several times more active than steel. But, if the test is continued, it gradually goes down to approximate equality with it. This was attributed to the action of active nitrogen on the metal surface, which, after the experiment, was found to be no longer bright, as it had been initially. Like the photoelectric effect of zinc, this action of copper requires a clean surface. The experiment was repeated several times, starting in each case with a newly cleaned copper surface, and continuing the run till it had become about equal to the steel, so that no important change was observed on commutation. Thin copper foil allowed to become red hot in active nitrogen is left clean with a red spongy surface like an electrolytic deposit (Rayleigh 1940, p. 17). The formation of a film, presumably nitride, occurs at a lower temperature, and clears off if the copper gets hot, leaving the surface spongy.

Another experiment tried with the same general arrangement was to compare a cathode of perforated zinc, backed closely by sheet glass, with a cathode of continuous zinc sheet. The area of the former was only about half that of the latter, but no appreciable difference was observed between them. This again tells against there being any important surface effect with zinc.

5. COMPARISON OF LIGHT QUANTA EMITTED WITH NUMBER OF IONS PRODUCED

It is evidently of importance to compare the number of light quanta emitted by the afterglow per cubic centimetre per second with the number of ion pairs emitted. In this way it can be determined whether the processes of light emission and of ionization are necessarily bound up together or not.

The method of experimenting was to observe the afterglow photoelectrically in a 5 l. globe, coated with metaphosphoric acid, and to measure the ionization in an annexe vessel with coaxial cylindrical electrodes (figure 4). These electrodes are of metaphosphoric acid smeared on glass, and each making contact to a platinum wire carried through. In this case the afterglow was generated in another globe altogether separate, and was admitted beforehand by the tap at the top, which was closed while the experiment was in progress. The glow took a time comparable with 1 hr. to decay over the range examined.

Preliminary experiments, which it is hardly worth while to describe in detail, showed that the intensity of the glow in the annexe remained in a constant ratio to the intensity in the main vessel, and approximately in the same ratio as the respective depths in the direction of vision. It is assumed, therefore, that the brightness (number of quanta emitted per cubic centimetre per second) is the same in the main vessel as in the annexe. This results from the free diffusion at low pressures. The reason for using an annexe rather than carrying out both measurements in the same vessel is that for viewing faint lights a large volume is desirable, while for measuring saturation currents satisfactorily we require the small volume and short electrode distance.

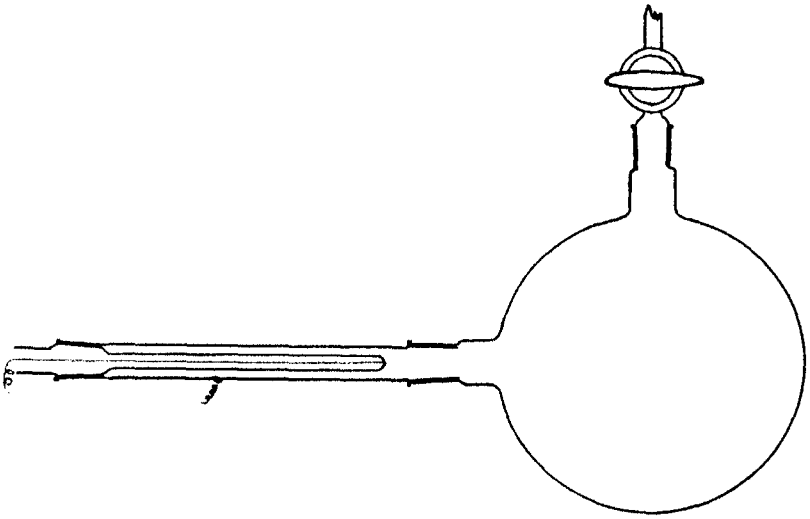


FIGURE 4 ($\frac{1}{8}$ actual size)

To determine the absolute luminosity in the bulb, it was necessary to calibrate the uranium photometer. This was done by observing with it a surface of magnesium oxide illuminated by light of known metre-candle value. A standard lamp was used of 0·735 c.p. at 19 m. Glass no. 12 equalized the lights. According to these data the brightness of the magnesium oxide is $6{\cdot}0 \times 10^{-8}$ candles per sq. cm., and of the uranium $3{\cdot}79 \times 10^{-9}$ candles per sq. cm. It was with the latter that the afterglow bulb was compared. The thickness looked through was 23 cm. When this matched the uranium, the brightness of the gas was $1{\cdot}65 \times 10^{-10}$ candles per c.c. Assuming that the afterglow light is concentrated at the most luminous part of the spectrum (λ 5560) (on this point see Rayleigh 1940, p. 2), this represents $1{\cdot}65 \times 10^{-10} \times 2{\cdot}02 \times 10^{-5}$ or $3{\cdot}33 \times 10^{-5}$ erg/sec./c.c. The quantum at this frequency is $3{\cdot}57 \times 10^{-12}$ erg. Accordingly the above emission amounts to $3{\cdot}33 \times 10^{-5}/3{\cdot}57 \times 10^{-12}$, or $9{\cdot}33 \times 10^{6}$ quanta/c.c./sec.

The saturation current between the inner and outer cylinders was determined by using as before a Lindemann electrometer with an alcohol xylol resistance of $1{\cdot}1 \times 10^{9}$ ohms, through which the charge could leak away to earth. This resistance had been determined by a direct deflection method, applying 100 V. 1 division deflexion of the electrometer represented $1{\cdot}39 \times 10^{-11}$ amp. It was necessary to pay constant attention to the position of the zero, which was determined after each reading. The volume of the space between the electrodes was 42·5 c.c.

The calculation of an experiment may be given. The intensity was 126 times the uranium unit. This represents $126 \times 9·33 \times 10^{6}$ or $1·18 \times 10^{9}$ quanta/c.c./sec.

The current was $62·6 \times 10^{-11}$ amp., or $1·47 \times 10^{-11}$ amp./c.c. of gas space.

Taking the charge on an ion as $1·59 \times 10^{-19}$ coulombs, this means $1·47 \times 10^{-11}/1·59 \times 10^{-19}$ or $9·26 \times 10^{7}$ pairs of ions/c.c./sec.

Thus the number of quanta/number of ion pairs $1·18 \times 10^{9}/9·26 \times 10^{7}$ or 13 quanta for every pair of ions produced.

With this example the series of results may now be given.

The glow was allowed to decay over a wide range, one observer noticing when it passed a certain brightness, and another taking the current reading at the same moment. The following are the means of two very concordant runs:

intensity of light (in terms of uranium standard)	126	63·1	31·6	15·8	7·94	3·98	2·00	1·00
current, amp. $\times 10^{-11}$	62·6	38·9	26·8	18·9	12·9	9·17	7·09	5·91
number of quanta/ number of ion pairs	13	10	7·4	5·2	3·9	2·7	1·8	1·1

It appears therefore that the number of quanta at first exceeds the number of pairs of ions generated by a considerable multiple, but that as the glow decays this multiple becomes less, and towards the end of the set of observations it does not differ much from unity. Had it been possible to deal with smaller intensities and currents a stage would probably have been reached when the number of ion pairs would have exceeded the number of quanta.

Reviewing these results, the observed facts may be expressed by saying that the ionization decays much more slowly than the luminosity, and consequently gains in relative importance as the decay proceeds. Over the extreme range the light varies as 126 to 1. The square root of this ratio is 11·2. The ionization varies in the ratio 10·6 to 1. Thus it would not be far from the truth to say that the ionization is as the square root of the luminosity, though I consider that this law is suggested rather than proved.

If the ionization were part of the normal process involved in production of the glow, it might be expected that these two processes would proceed *pari passu*. For example, it might be thought likely that one pair of ions would be produced for every light quantum that was emitted; or, if not this, that some other small multiple or submultiple would be maintained. Although during part of the decay process the ionization is of this order of magnitude, yet nothing like an invariable ratio is preserved, the ratio quanta/ion pairs changing from 13 to 1·1 over the observed range of decay.

It is concluded that the process leading to ionization is not the same as that leading to light emission. It has been shown (Rayleigh 1940, pp. 8–10) that the light emitted is as the square of the concentration of the active material, if the concentration of inert nitrogen is constant.

If we take it that the ionization is as the square root of the light emission and combine these results, it appears that the ionization is as the first power of the concentration of active material, other conditions being constant.

6. RESULT OF ADDING INERT NITROGEN

It was shown, originally by Kneser (1929) and later by myself (1940), that the nitrogen afterglow was temporarily increased in intensity if, in the course of its decay, inert nitrogen is added. It was of considerable interest to determine whether or not this would produce an increase in the ionization. For this experiment a 5 l. globe was used coated internally with metaphosphoric acid. An electrode, also of glass coated with metaphosphoric acid, protruded into it. This was connected to the electrometer used with a high resistance, and an e.m.f. of 9 V was applied to the outer coating.

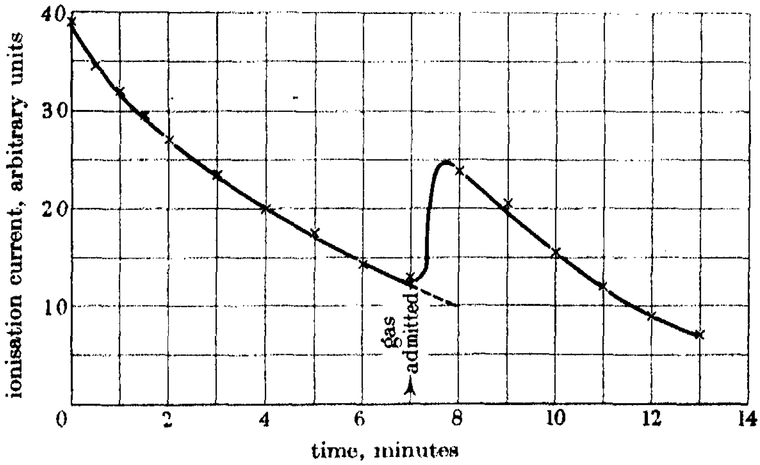


FIGURE 5

The bulb was filled with glowing nitrogen at 4×10^{-3} cm. pressure, and the generating vessel shut off by a stopcock. When the intensity had decayed to a convenient point readings were started. They are shown in figure 5. After 7 min. the pressure was increased from 4×10^{-3} to $10{\cdot}5 \times 10^{-3}$ cm., i.e. 2·6 times. At 8 min. the current was 24 arbitrary units, and extrapolating from its previous trend it is estimated that its value would then have been 10 units if the addition of gas had not been made. Thus the increase of current is 2·4 times, i.e. in about the same ratio as the increase of pressure. It is quite certain therefore that the ionization is increased by increase of inert nitrogen concentration, and the indication is that it is also increased in the same ratio.

To sum up it seems probable that the ionization is directly as the concentration of the active material, the concentration of inert nitrogen being constant; and also that it is directly as the concentration of inert nitrogen, the concentration of active material being constant. The natural conclusion would be that ionization results from single collisions between active centres and nitrogen molecules.

The afterglow intensity, on the other hand, is as the square of the concentration of the active centres, the inert nitrogen concentration being held constant: while in respect of inert nitrogen concentration it behaves like the ionization, that is, it varies as the first power of the concentration of inert nitrogen, if the concentration of the active material is not varied (Rayleigh 1940, pp. 6–8).

The ionization process and the afterglow process are thus considered in some measure independent. Nevertheless, the ionization effect has an important bearing on the question of the energy of active nitrogen. The ionization potential of the nitrogen molecule is 15·51 V, and unless some process of multiple excitation is involved, the active centres which are in collision with nitrogen molecules must have this amount of energy, or something very near it in order to produce ionization. This is much greater than the energy of 9·6 V attributed to active nitrogen on spectrographic grounds (Saha and Mathur 1936). It seems to me that those who have discussed the active nitrogen problem mainly from the spectroscopic point of view have ignored the ionization phenomenon too much. It may have been supposed that this was entirely subordinate, but the present work shows that the ionization may under some conditions involve as many atoms as the light emission.

I have received valuable help from my assistant, Mr R. Thompson, in these investigations.

REFERENCES

Constantinides, P. A. 1927 *Phys. Rev.* **30**, 95–108.
Kneser, H. O. 1929 *Ergebn. Exakt. Naturw.* **8**, 229–257.
Rayleigh 1940 *Proc. Roy. Soc.* A, **176**, 1–27.
Saha, M. N. and Mathur, L. S. 1936 *Proc. Nat. Acad. Sci. India*, **6**, 120–127.
Strutt, R. J. (Rayleigh) 1911 *Proc. Roy. Soc.* A, **86**, 56–63.
Strutt, R. J. (Rayleigh) 1912 *Proc. Roy. Soc.* A, **87**, 179–188.

Surface energy relations in liquid/solid systems

I. The adhesion of liquids to solids and a new method of determining the surface tension of liquids

BY G. MACDOUGALL, PH.D. AND C. OCKRENT, D.SC.

(*Communicated by J. Kendall, F.R.S.—*
Received 28 *March* 1941.—*Revised* 10 *September* 1941)

A new method for determining the surface tension of liquids has been derived. This involves the consideration of the advancing and receding contact angles of a liquid drop on a tilted solid surface. The theory has been tested by an improved optical projection technique for a variety of liquid/solid systems and the results obtained are in agreement with the accepted values. It is shown that the advancing and receding contact angles are characteristic constants of liquid/solid systems and the calculated and measured values of the minimum receding angles are in agreement. The prevailing views of 'hysteresis' effects or 'stationary' contact angles which have arisen to account for the data available are incorrect and the discordant experimental results reported are due to inadequate technique. The difference between the adhesions corresponding to the advancing and receding angles is ascribed to the work done in removing an adsorbed unimolecular layer. The work done in gcal./mol. in forming this adsorbed layer is in reasonable agreement with that expected from studies in gas/solid systems and the forces involved are van der Waals'. Further, different solids that might be expected to show similar surface structures yield similar values for the work done. The variation in the value of the advancing angle in some liquid/solid systems and its constancy in others is reconciled with the polar character of the solid surface, i.e. it is suggested that short-range forces are involved. It has been found that monolayers of ferric stearate on glass are orientated with their hydrocarbon tails away from the interface in agreement with electron diffraction measurements. It is suggested that the methods may be useful for investigating the structure of monofilms and built-up layers of monofilms.

INTRODUCTION

The wetting of solids by liquids has been extensively studied in recent years probably because many technical processes depend upon wetting. The work of adhesion, i.e. work done in separating unit area of liquid from the solid, was given by Dupré (1869) as

$$W_{SL} = \nu_{SA} + \nu_{LA} - \nu_{SL},$$

where W_{SL} = work of adhesion of liquid for the solid,
ν_{SA} = the surface tension of the solid,
ν_{LA} = the surface tension of the liquid,
ν_{SL} = interfacial tension at solid/liquid interface.

This cannot be evaluated directly since no satisfactory method exists for the direct determination of the surface tensions of solids or of the interfacial tensions of solids against liquids. Young (1805) expressed the work of adhesion in terms of the horizontal components of the surface tensions as

$$W_{SL} = \nu_{LA}(1+\cos\theta),$$

where θ is the contact angle measured in the liquid. For liquids which completely wet the solid, $\theta = 0$ and $W_{SL} \geqslant 2\nu_{LA}$.

For liquids of known surface tensions the work of adhesion could be found if there was a single value for the contact angle. In practice many discordant values for the same system have been recorded. In general, different values are obtained according to whether the liquid is advancing or receding over the surface.

These differences have been described by Sulman (1920) and by Edser (1922) as 'hysteresis' effects. Pockels (1914), Bartell (1933), Adam (1938) and others have suggested that they may be due to contamination of the solid or liquid surfaces. Ablett (1923), using the rotating cylinder machine method, obtained consistent values for the advancing and receding angles of water on paraffin wax and suggested the difference may be due to adsorption of water by the wax.

Generally, it is suggested that the correct contact angle is the mean of the advancing and receding angles and this is called the 'equilibrium contact angle'.

THEORETICAL

Equilibrium of a drop on a tilted plate

Consider the equilibrium forces on a drop resting on an inclined plate (figure 1).

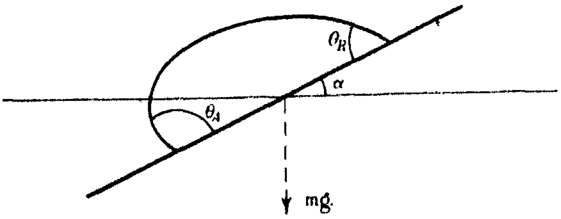


FIGURE 1

Let α be the angle of inclination of the plate. The contact angle, measured in the liquid, of the lower edge of the drop is θ_A (advancing angle) and the contact angle of the upper edge is θ_R (receding angle). ν and ρ are the surface tension and density of the liquid respectively.

Now take a lamina of mass m, thickness dl, and area A through the centre of gravity of the drop, at right angles to the plane surface and to the axis of inclination.

Since the drop is symmetrical about the lamina, at equilibrium the forces acting on it can be resolved along the plate in a direction parallel to the lamina.

Therefore
$$\begin{aligned}\nu dl\cos\theta_A &= \nu dl\cos\theta_R - mg\sin\alpha\\ &= \nu dl\cos\theta_R - A\rho g\,dl\sin\alpha,\end{aligned}$$
since $m = Adl\rho$, and
$$\nu(\cos\theta_R - \cos\theta_A) = A\rho g\sin\alpha, \tag{1}$$
or
$$\cos\theta_R - \cos\theta_A = K_1 A\sin\alpha, \tag{2}$$
where
$$K_1 = \frac{\rho g}{\nu}.$$

If θ_A, θ_R, A and α are measured and $(\cos\theta_R - \cos\theta_A)$ plotted against $A\sin\alpha$, a straight line should result from whose slope the surface tension can be calculated.

Advancing angle constant

Consider the special case where the advancing angle may remain constant on tilting the plate; (2) then becomes
$$\cos\theta_R = K_1 A\sin\alpha + K_2, \tag{3}$$
where $K_2 = \cos\theta_A$(const.).

Here the curve of $\cos\theta_R$ against $A\sin\alpha$ will be linear and cut the $\cos\theta_R$ axis at a point corresponding to K_2.

Minimum receding angle

The receding angle of a drop resting on a plate diminishes as the plate is tilted. An angle of tilt may be obtained at which the drop begins to slide. If the receding angle at this point reaches a minimum and is a characteristic constant, then from (2)
$$K_1 A\sin\alpha = K_3 - \cos\theta_A, \tag{4}$$
where $K_3 = \cos\theta_R$(min.).

Further, if the advancing angle is constant at any angle of tilt or becomes constant before or at the angle of tilt at which the drop begins to slide, then (4) reduces to
$$A = \frac{K_4}{\sin\alpha}, \tag{5}$$
where
$$K_4 = \frac{\cos\theta_R(\text{min.}) - \cos\theta_A(\text{const.})}{\rho g/\nu}.$$

From (5), A is inversely proportional to $\sin\alpha$, where α is the value of the angle of tilt at which sliding first occurs. If A, θ_A(const.) and α are measured, then θ_R(min.) can be calculated from the slope. (Hereinafter the sliding angle of tilt will be referred to as α_s.)

The validity of (2), (3) and (5) can be tested as follows: Equations (2) and (3) can be tested by using drops of any size, and (5) by the behaviour of a series of drops of different size. From (2), which includes the special case of (3), the surface tension of the liquid concerned should be calculable. From (5) the minimum receding angle should be calculable.

EXPERIMENTAL

Apparatus

A drop of liquid was placed on the surface of a thin flat plate of the solid (*C*, figure 2) and illuminated by a parallel beam cooled by the water trough (*B*).

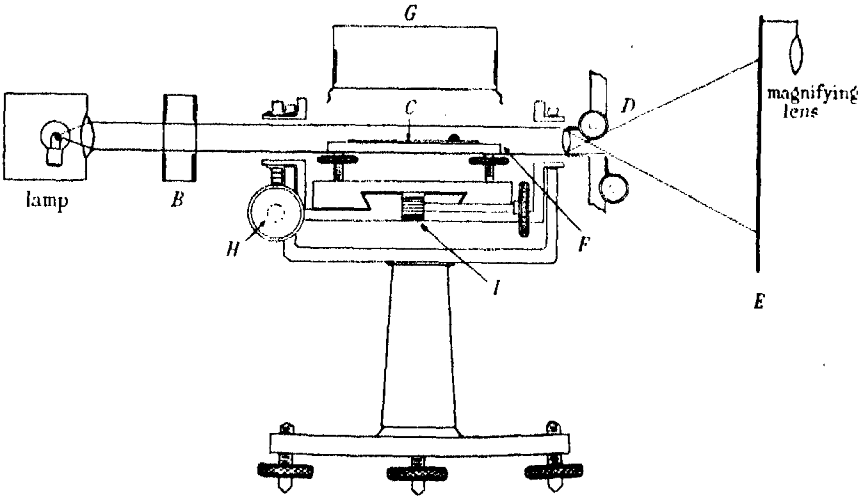


FIGURE 2

The travelling microscope (*D*) had vertical and horizontal adjustments for the small corrections often required after tilting the stage. The microscope eyepiece was removed and the image of the drop thrown on the screen (*E*) by a 1·5 in. achromatic objective. The magnification was about 18 diameters. By placing a fine hair line, photographed on a glass slip, tangentially to the point of contact of the liquid and solid the angle of contact could be read off directly; the reproducibility was $\leqslant 0{\cdot}5^\circ$.

The solid surfaces were thin films on glass microscope slides mounted on the stage (*F*). This was covered with a close-fitting metal lid (*G*) provided with two glass windows for illuminating the drop. The effect of tilting on the contact angles was studied by mounting the stage on large annular bearings operated by a worm gear (*H*) enabling the stage to be rotated through 90° on either side of the horizontal. The bearings were of sufficient diameter to accommodate the illuminating beam on one side and the microscope objective on the other. As the axis of rotation of the stage was collinear with the axis of the microscope, the image of a point object (contact of liquid and solid) focused on the cross-wires of the screen remained stationary on tilting the stage. This makes for speed and convenience of measurement. The stage was supported on four spring-loaded screws so that the plane of the specimen could be raised or lowered into the axis of rotation, and provided with a rack and pinion (*I*) so that the drop could be traversed across the microscope field and both ends of the outline observed.

The whole apparatus, except for the lamp and screen, was mounted on a massive lithographic stone supported by four rubber bungs. This diminished vibration and restored the system rapidly after any disturbance.

The screen was prepared by photographing a scale on a half-plate and precipitating barium sulphate in the emulsion by alternate immersion in barium chloride and ammonium sulphate. This gave a sharper image, freer from colour, than ground glass. The screen was mounted in a metal frame which allowed it to be raised or lowered and tilted and the image on the screen was viewed through a magnifying lens.

Materials

The water used was distilled from alkaline permanganate into a pyrex apparatus from which, by a pressure device, drops of water from the interior of the bulk were ejected through a capillary tube drawn out to a fine jet, the tip of which was ground flat. The water was free from surface contamination for it gave reproducible contact angles on various surfaces over a period of many months.

The mercury was washed in dilute nitric acid and distilled twice in vacuo.

Solid surfaces

The stearic acid, m.p. 67·5° C, was prepared from sodium stearate and crystallized from ether, and the cetyl palmitate, m.p. 48·2° C, from spermaceti by crystallization from alcohol. Commercial cetyl alcohol, m.p. 49·2° C, was crystallized from alcohol; a second specimen prepared from cetyl palmitate and crystallized from alcohol had a m.p. of 48·7° C. Both samples gave the same contact angle. The tristearin, an impure sample, m.p. 59° C, was used without further purification. Ferric stearate, m.p. 91° C, and the rubbed down monofilms on glass were prepared by the method of Blodgett and Langmuir (1937).

The glass microslides were roughened with carborundum powder and after treatment with chromic acid cleaning mixture were washed with the water prepared as above and dried in an air oven. The criterion for the cleanliness of the slides was the zero contact angle obtained with water.

At the beginning difficulty was experienced in obtaining reproducible contact angles. This was traced to rough uneven films. Even, reproducible films were finally obtained by drawing a clean glass straight edge across the ground microscope slide which carried a thin film of the molten material.

Measurement of cross-sectional areas of drops

The drop image was thrown on squared paper placed against the screen and sufficient points marked off on its contour for the complete outline to be drawn later. The squares were counted and the area computed from the magnification of the optical system. Reproducible results were obtained for medium sized drops to < 1 %. Where evaporation of the drop does not occur as in mercury, or is prevented as with water by humidifying the atmosphere of the drop, the cross-sectional area was found to remain sensibly constant on tilting the plate. In cases where the areas varied during the run, they were measured after each increment of tilt.

All the measurements described here were obtained in a room maintained at 18° C ± 2.

RESULTS

(*a*) *Water on solid surfaces*

The curves of ($\cos\theta_R - \cos\theta_A$) against $A\sin\alpha$ were obtained for water on the following surfaces:

(1) paraffin wax, (2) cetyl palmitate,
(3) tristearin, (4) stearic acid,
(5) ferric stearate (monofilm on glass).

Figure 3 shows the results for stearic acid and tristearin, and the points for both systems lie on the same straight line. The curves for paraffin wax, cetyl palmitate and ferric stearate (monofilm) also coincide with this line, but their points are not shown since they tend to be crowded near the bottom when placed on the same scale; the measurements of θ_A, θ_R, and α are, however, given in table 1.

The results are as predicted by equation (2) and the curve is common for the different surfaces and thus independent of the type of solid surface, the initial contact angle and the size of the drops and only dependent on the nature of the liquid. The surface tensions, calculated from the slopes of the individual curves, are recorded in table 2. They are in reasonable agreement with each other and with the accepted values obtained by other methods. The slightly lower values on ferric stearate monofilms are probably due to the method of preparation of these films (later).

The results quoted in table 2 are representative of many different concordant runs.

On paraffin wax, cetyl palmitate and ferric stearate (monofilm) θ_A is constant and independent of α, so in these cases equation (3) also applies (table 1).

In practice in all cases where θ_A does not vary with α it is only necessary to measure θ_R as the plate is tilted in order to obtain ν.

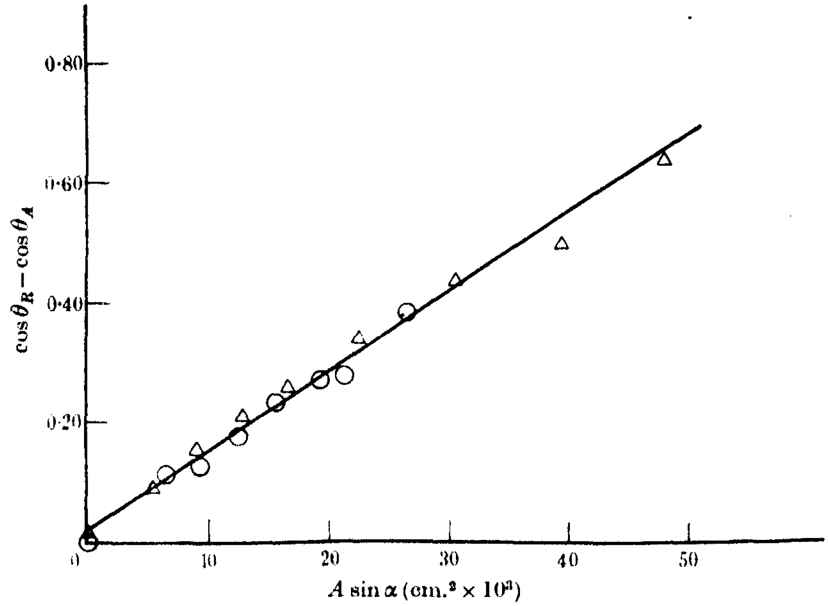


FIGURE 3. ○○ water on stearic acid; △△ water on tristearin.

TABLE 1

water on paraffin wax			water on cetyl palmitate			water on a monofilm of ferric stearate		
$A = 0{\cdot}0610$ cm.^{2}			$A = 0{\cdot}102$ cm.^{2}			$A = 0{\cdot}0685$ cm.^{2}		
θ_A	θ_R	α	θ_A	θ_R	α	θ_A	θ_R	α
111·0°	111·2°	0°	109·8°	109·9°	0°	93·0°	93·0°	0°
111·5°	109·7°	1·2°	110·0°	108·5°	1·0°	93·2°	91·2°	1·6°
111·1°	108·5°	3·0°	110·6°	106·7°	2·4°	93·4°	89·2°	3·3°
111·6°	107·5°	4·2°	110·6°	105·7°	3·1°	93·8°	88·3°	5·2°
111·2°	106·3°	5·8°	110·5°	102·9°	5·2°	93·9°	86·7°	6·8°
111·5°	104·1°	8·6°				93·5°	85·3°	8·1°
111·5°	102·2°	10·2°				93·6°	83·9°	10·0°

(*b*) *Mercury on solid surfaces*

Figure 4 shows that the curve of ($\cos\theta_R - \cos\theta_A$) against $A \sin\alpha$ for mercury on glass, paraffin wax and cetyl alcohol is linear and common for all and the results are as predicted by (2). The surface tension of mercury calculated from the slopes of the separate curves is shown in table 2. The values agree with each other and with those obtained by other methods.

In this series, only for cetyl alcohol did θ_A increase on tilting until it reached constancy. On paraffin wax and glass θ_A was constant at all angles of tilt. Here therefore equation (3) also applies.

TABLE 2

liquid	solid	ν_A (ergs/cm.^{2}), calculated from curves 18° C ± 2		ν_A (ergs/cm.^{2}) (other methods)	
water	paraffin wax	73·0	± 1·5	73·05 at 18° C	
	cetyl palmitate	72·5		(*Int. Crit. Tables*)	
	tristearin	71·5			
	stearic acid	71·2			
	ferric stearate (monofilm on glass)	69			
mercury	glass	480	± 15	48_5 at 20° C	
	paraffin wax	470		(Adam 1938)	
	cetyl alcohol	485			
linseed oil	paraffin wax	32·0	± 1·5	32·5–33	
	manucol	31·7		(various sources)	
	gum arabic	33·5			
0·2 N sodium hydroxide	paraffin wax	72·0	± 1·5	73·4	18° C
1·0 N sodium hydroxide	paraffin wax	73·5		74·9	
				(*Int. Crit. Tables*)	

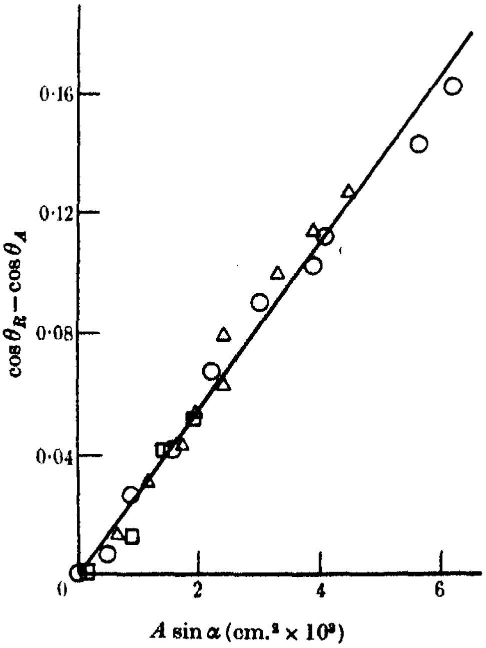


FIGURE 4. ○○ mercury on glass; △△ mercury on paraffin wax; □□ mercury on cetyl alcohol.

(c) *Linseed oil on various surfaces*

Drops of linseed oil have been studied on paraffin wax, gum arabic and manucol (sodium alginate). These combinations are of technical interest in printing. Figure 5 shows that $(\cos\theta_R - \cos\theta_A)$ is the same linear function of $A \sin\alpha$ for all three surfaces in accord with equation (2). The surface tension values for the oil (calculated from the separate curves) are in agreement with each other and the accepted values (table 2). θ_A was independent of α for paraffin wax and so here equation (3) also applies.

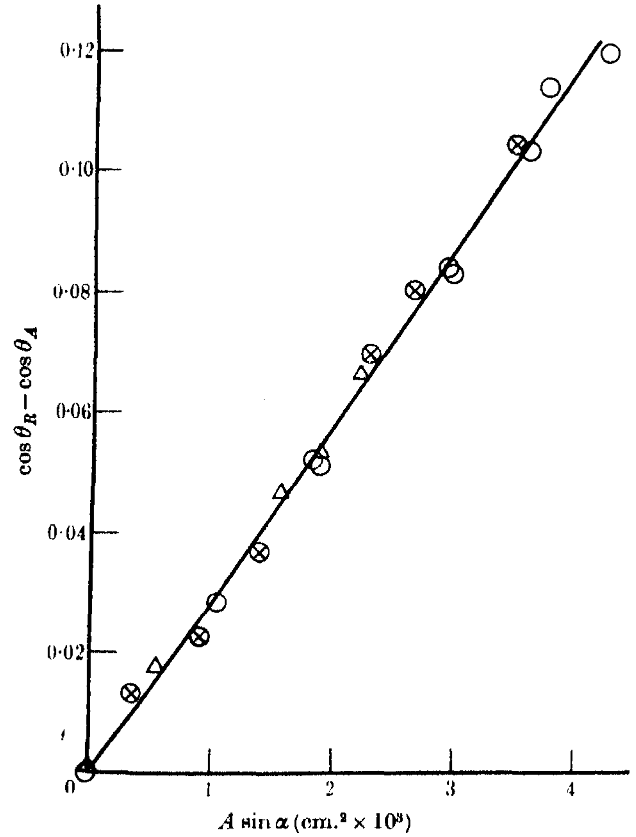


FIGURE 5. ○○ linseed oil on paraffin wax; △△ linseed oil on gum arabic; ⊗⊗ linseed oil on manucol.

(d) *Solutions of sodium hydroxide on paraffin wax surfaces*

Figure 6 shows that $(\cos\theta_R - \cos\theta_A)$ is a linear function of $A \sin\alpha$ for 0·2N and 1·0N sodium hydroxide solutions. The calculated surface tensions (table 2) also agree reasonably well with those from other methods. Here equation (3) also applies since θ_A is independent of α.

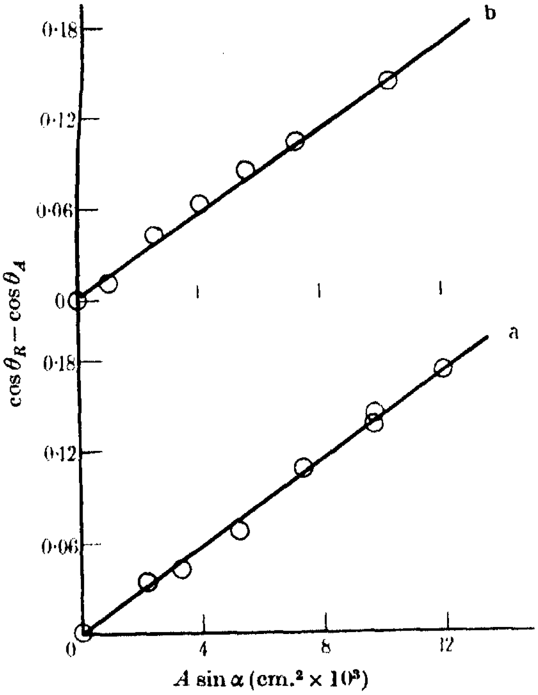


FIGURE 6. (b) 1·0 N sodium hydroxide on paraffin wax; (a) 0·2 N sodium hydroxide on paraffin wax.

Partially evaporated drops

A water drop on a horizontal paraffin wax surface comes to rest with an angle of contact $=\theta_A$(const.). On evaporation this angle decreases, the drop remaining symmetrical.

To find if appreciable evaporation affects the relationship established between ($\cos\theta_R - \cos\theta_A$) and $A \sin\alpha$, drops of water on paraffin wax were evaporated until the contact angle was less than θ_A(const.) = 111·5°. Figure 7 shows the composite curve for two drops whose initial angles decreased to 108° and 106° respectively. In accord with equation (2), the relationship is linear and independent of the initial advancing angle. The surface tension of water calculated from the slope (72·4 ± 1·5 ergs/cm.^{2}) agrees with the accepted value.

If a drop partly evaporates on a horizontal surface and is subsequently tilted, the initial contact angles alter. The angle of the lower edge increases while the hindmost angle approaches the value of the minimum receding angle (see later). At α_s, when the drop begins to move, it has the same advancing (θ_A const.) and receding (θ_R min.) angles of initially normal unevaporated drops at their respective values of α_s.

The advancing angle

For systems where the advancing angle is constant (θ_A independent of α), disturbing influences such as evaporation, vibration, etc. must be excluded to obtain reproducibility.

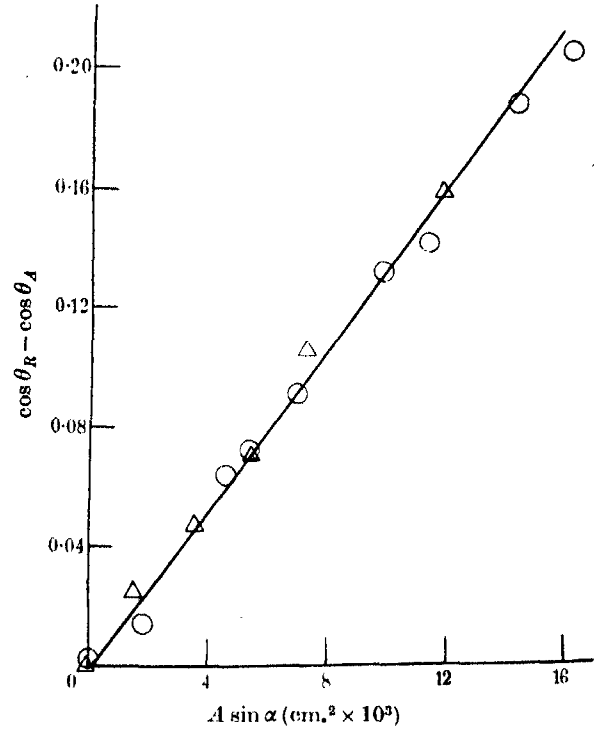


FIGURE 7. Partially evaporated drops; water on paraffin wax. ○○ initial $\theta_A = 106{\cdot}2°$; △△ initial $\theta_A = 108{\cdot}0°$.

In table 3 are listed those surfaces on which θ_A is constant and those on which θ_A increases with α. In table 4 are given the values of θ_A(const.) for all the systems investigated.

TABLE 3

liquid	θ_A constant	θ_A increases
water	paraffin wax	tristearin
	cetyl palmitate	stearic acid
	ferric stearate (monofilm)	—
mercury	paraffin wax	cetyl alcohol
	cetyl palmitate	stearic acid
	glass	—
linseed oil	paraffin wax	gum arabic
	—	manucol
sodium hydroxide (0·2 N)	paraffin wax	—
sodium hydroxide (1·0 N)	paraffin wax	—

The variation of θ_A and its relation to the type of surface concerned is discussed later.

TABLE 4

			θ_R(min.)		W_A (ergs/cm.^{2})	W_R (ergs/cm.^{2})	w (ergs/cm.^{2})	H (gcal./gmol.)
liquid	surface	θ_A(const.)	calc.	obs.				
water	paraffin wax	111·5°	96·0°	96°	45·9	64·9	9·5	163
	cetyl palmitate	110·3	96·6	96·5	47·3	64·1	8·4	145
	stearic acid	105	53	55	53·7	116	31·2	538
	cetyl alcohol	25	—	0	138	—	—	—
mercury	glass	139·0	114·5	118	119	284	83	1134
	paraffin wax	149·0	135·5	137	69·2	139	35	483
	cetyl palmitate	149·8	135·0	134	65·2	149	42	580
	stearic acid	150·0	—	—	45·1	—	—	—
	cetyl alcohol	153·5	—	—	49·5	—	—	—
linseed oil	gum arabic	28	—	0	61·3	—	—	—
	manucol	39	—	0	57·8	—	—	—
	paraffin wax	45	—	0	55·5	—	—	—
1·0N NaOH	paraffin wax	110·0	95·7	—	48·4	66·2	8·9	153
0·2N NaOH	paraffin wax	109·5	95·6	—	48·6	65·8	8·6	148

Even though θ_A may not have the value θ_A(const.) before tilting, it is found, in all the cases completely investigated, it does attain a constant reproducible value at or before α_s is reached. This is illustrated by examples in table 5.

TABLE 5. ADVANCING ANGLE REACHING CONSTANCY

water on stearic acid		linseed oil on manucol	
θ_A	α	θ_A	α
102·0°	0°	35·6°	0°
103·2°	6·0°	36·3°	1·2°
104·1°	8·1°	37·0°	3·1°
104·5°	10·3°	37·3°	4·8°
105·4°	13·2°	37·9°	6·1°
105·2°	15·0°	38·0°	7·9°
		38·7°	9·2°
		38·9°	10·9°

mercury on stearic acid		mercury on tristearin	
θ_A	α	θ_A	α
146·0°	0°	148·3°	1·7°
149·3°	9·7°	160·0°	8·0°
155·3°	16·2°	159·9°	12·1°
155·0°	20·4°		

The minimum receding contact angle

From equation (5) θ_R(min.) should be calculable from the slope of A against $1/\sin\alpha_s$, this applies when θ_A is constant and does not change with α or when it attains a characteristic constant value, θ_A(const.), for a given system at or before α_s. The value of θ_A(const.) must therefore be known to calculate θ_R(min.).

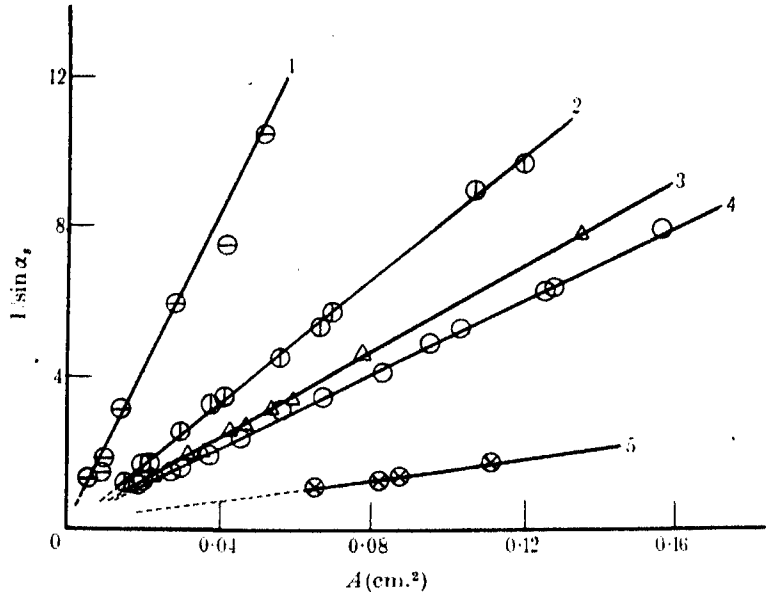


FIGURE 8

Figure 8 shows the curves of A plotted against $1/\sin\alpha_s$, using drops of different sizes, for the following systems:

(1) mercury on paraffin wax,	(2) mercury on glass,
(3) water on cetyl palmitate,	(4) water on paraffin wax,
(5) water on stearic acid.	

The curve for mercury on cetyl palmitate is almost coincidental with that of paraffin wax and is therefore not shown.

Figure 9 shows the corresponding curves for 0·2N and 1·0N sodium hydroxide on paraffin wax.

For water on stearic acid θ_A becomes constant before α_s is reached. For the other cases θ_A is independent of α.

As predicted by equation (5) figures 8 and 9 show that A is a linear function of $1/\sin\alpha_s$. The values of θ_R(min.) calculated from the slopes are given in table 4. It is difficult to verify the result experimentally since the receding angle of the drop cannot be measured just as it begins to move off the plate. An approximate value can however be obtained by taking frequent measurements of the receding angle at angles of tilt as close as possible to α_s and also by other methods (later). Where this has been done the agreement is reasonably satisfactory (table 4).

The significance of the characteristic advancing and receding angles is considered in the discussion. In table 4 are also given the works of adhesion W_A and W_R calculated from θ_A(const.) and θ_R(min.) respectively.

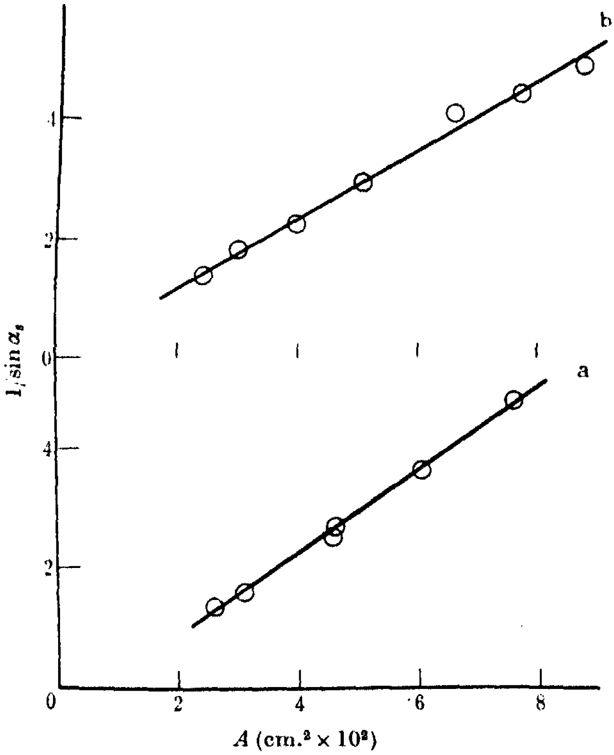


FIGURE 9. (b) 1·0N sodium hydroxide on paraffin wax; (a) 0·2N sodium hydroxide on paraffin wax.

Discontinuities in the contraction of drops and the minimum receding angle. A direct observation of the minimum angle is obtained by the following method. A drop of water was placed on paraffin wax. It was observed that the contact angle decreased steadily on evaporation, without any diminution of the liquid/solid contact area, from its initial value of 111·5° to about 104° when the drop suddenly contracted and the contact angle rose to 110°. As evaporation proceeded this cycle was repeated many times, the contact angle at which contraction occurred diminishing with each cycle until it reached 96°, i.e. the value of the minimum receding angle (table 4). The drop then evaporated regularly without further change in the contact angle, the liquid/solid contact area diminishing steadily.

The contact angle at which a drop falls off. The receding angle can also be observed directly as the drop is actually moving as follows. A glass tube was plugged at one end with paraffin wax. Through a fine hole drilled in the wax water was allowed to drip slowly. At the outlet a drop gradually formed and spread across the wax surface and, when large enough, the edge of the drop receded over the surface as the drop became elongated and began to detach itself from the wax. It is difficult to measure accurately the contact angle as the drop is nearly falling off but it lay between 95–100°, i.e. in agreement with the calculated value of 96°.

DISCUSSION

Figures 3–7 are in accord with equations (2) and (3) and are representative of a wide variety of systems. Also the surface tensions calculated from the slopes (table 2) agree well with the accepted values for pure liquids (water and mercury), a mixture (linseed oil) and for aqueous solutions. Tables 1 and 5 show the advancing angles are constant and reproducible at all angles of α or become so. (Those systems where θ_A increases with α are discussed later.) $1/\sin\alpha$ is a linear function of A (figures 8, 9) as predicted by equation (5) and the values of θ_R(min.) calculated from the slopes are reproducible and independent of the size of the drop and θ_R(min.) is characteristic of any given liquid/solid system just as is θ_A(const.). Where it can be verified the calculated and observed values of θ_R(min.) agree reasonably well (table 4).

It may therefore be concluded that the theory advanced here is verified and that the prevailing conceptions of 'equilibrium contact angles' or 'stationary contact angles' are wrong. The so-called 'hysteresis' of the contact angles (i.e. difference between θ_A and θ_R) is not due to variable frictional forces or sorption of liquid into cracks of the solid (Edser 1922). On the contrary, it can be stated that a liquid drop when $\theta_A \neq 0$ exhibits two characteristic, constant and reproducible contact angles θ_A(const.) and θ_R(min.).

The theory and method described here could provide a useful addition to those at present available for measuring surface tensions. The accuracy could be improved by further refinements.

The relationships elucidated can only be established by scrupulous cleanliness and hold for all combinations of solids and liquids, if the following criteria are satisfied. The liquid and solid surfaces must be uncontaminated by extraneous impurity and the liquid and solid must be mutually insoluble and not react chemically. In the case of mixed solids or liquids their constituents must also be undisturbed by solubility effects.

Confusion has been caused by failure to appreciate the results of cleansing experiments in previous work. Edser (1922) noted that glass covered with a greasy film gave finite contact angles and that a 'hysteresis' was observed in the contact angle depending whether the liquid was advancing or receding. On removal of the grease by acid zero contact angle for water (i.e. complete wetting) was obtained. Similar effects were observed on galena, blende, etc. by Sulman (1920) and others. From this investigators concluded that 'hysteresis' is due to lack of cleanliness or inhomogeneity of the surface. The true explanation in these cases is that the 'hysteresis' is due to advancing and receding angles *characteristic of the greasy surface*, i.e. of a surface only partially wetted, where $\theta_A \neq 0$. Since a *clean* glass surface is completely wetted by water, $\theta_A = 0$ and hence the advancing and receding angles are equal.

Bartell and Wooley (1933) found that special cleaning of silica and pyrex gave approximately equal advancing and receding angles for α-bromnaphthalene. However, here the receding angles had very low values (5–0°) and so it appears that good cleaning resulted in complete (or almost complete) wetting and so $\theta_A \simeq \theta_R$, both tending to zero. This is supported by their own observations that on the same surfaces when, as the result of less effective cleaning, θ_R was larger θ_A was always $> \theta_R$.

Pockels (1914) found for a number of incompletely wetting liquids that $\theta_A = \theta_R$ on glass, platinum, copper and zinc. It appears, however, she was probably not really observing receding angles. The advancing angles were measured on the spread drops and the receding angles obtained by spreading the liquid drops with a wire and then measuring θ_R after the drops contracted to equilibrium. Under these conditions any angle between θ_A and θ_R is possible if the liquid is not actually moving or tending to move. Also her liquids had very low advancing angles, for example for benzene and carbon disulphide $\theta_A = 6°$ and 5·8° respectively, and the differences between θ_R and θ_A must have been small.

Often no clear distinction has been made between the 'hystereses' ascribed to the difference between θ_A and θ_R and the discordant results due to accidental contamination by an impurity. In the latter, if an impurity collects at the liquid/solid interface and W_{SL} differs for the contaminated and uncontaminated interfaces, the contact angle θ will alter. If the liquid/air interface is contaminated ν_{LA} will diminish and, even if W_{SL} is unaffected, then θ will change. In general impurities will affect ν_{LS} as well as ν_{LA} and so yield inconsistent contact angles. (An explanation for the difference between θ_A and θ_R is given later.)

Effect of chemical action on contact angle

The effect of chemical action, perhaps otherwise difficult to detect, is illustrated by the following experiment. The contact angles of a linseed oil drop, purified by molecular distillation,* were measured on zinc. The zinc was polished with emery, degreased with starch paste and afterwards with 2 % caustic soda solution and finally with a solution of alum in nitric acid (a lithographic preparing solution). The contact angle was found to change as follows:

---

* This sample was kindly supplied by Imperial Chemical Industries Ltd.

After 5 min. $\theta = 12°$. After 2 hr. $\theta = 3{\cdot}5°$.
After 10 min. $\theta = 9{\cdot}5°$. After 3 hr. $\theta < 2{\cdot}5°$.

The drop was left overnight and on removal a faint white deposit was observed.

Similar slow chemical or solubility effects may account for the discrepancies obtained by other investigators.

Reversibility

The change of contact angles with α is reversible. Thus for a mercury drop on paraffin wax a single linear curve of $(\cos\theta_R - \cos\theta_A)$ against $A\sin\alpha$ is obtained from measurements taken when the plate is tilted forward and backward in any order. The surface tension of mercury calculated from the slope is 465 ± 15 ergs/cm.

The significance of advancing and receding angles

Another explanation advanced for the difference between θ_A and θ_R is that the adhesion of a liquid drop is different for the unwetted and wetted surfaces and therefore θ_A, measured at the edge moving over unwetted solid, is $> \theta_R$, measured at the hindmost edge of the drop and thus moving over a portion of the surface already wetted by the liquid. This follows for $\cos\theta_R > \cos\theta_A$ if the adhesion of the liquid for wetted solid > unwetted solid (Young-Dupré).

To find if θ_R is merely characteristic of a previously wetted surface the following experiments were carried out.

A. A paraffin waxed slide was soaked in water and then quickly transferred to the highly humidified apparatus. A drop of water on the plate gave contact angles of 111·0° (for a clean, dry wax surface $\theta_A(\text{const.}) = 111{\cdot}5°$).

B. Large drops of water were placed on a waxed surface for different periods and when removed were quickly replaced within the same area by much smaller drops. The paraffin wax surface was maintained throughout at a very high humidity. The results obtained were as follows:

period large drop rested on wax min.	θ (small drops)
10	110·0°
22	109·2°
33	109·0°
65	109·0°

All this shows that preliminary soaking has no marked influence on the contact angle. The decreases are probably due to slight contamination on standing or in opening and closing the cover, for the contact angle is only slightly $<\theta_A$(const.) and much $>\theta_R$(min.) Adam and Jessop (1925) found that prolonged soaking in water of organic materials, varnished and painted surfaces and paraffin wax, gave decreases in the contact angles without abolishing the 'hysteresis'.

A more precise explanation for the difference between θ_A and θ_R when a liquid adheres to a solid (i.e. $\theta_A > 0$) is the following.

For equilibrium $$\nu_{SL} + \nu_{LA} \cos\theta_A - \nu_{SA} = 0. \tag{6}$$

If adhesion involves an adsorbed, orientated uni-layer of liquid molecules on the solid surface, then work will be done in forming this layer and (6) will be modified to

$$\nu_{SL} + \nu_{LA} \cos\theta_A - \nu_{SA} - w = 0, \tag{7}$$

where $-w$ is the work gained in forming the adsorbed layer. Similarly for the liquid receding (contact angle θ_R) the work gained will be $+w$, and

$$\nu_{SL} + \nu_{LA} \cos\theta_R - \nu_{SA} + w = 0. \tag{8}$$

From (7) and (8) $$\nu_{LA}(\cos\theta_R - \cos\theta_A) + 2w = 0, \tag{9}$$

whence $$-w = \frac{\nu_{LA}}{2}(\cos\theta_R - \cos\theta_A), \tag{10}$$

and this from Young's equation can be written as

$$-w = \frac{W_R - W_A}{2}. \tag{11}$$

In table 4, $-w$ is given for a number of systems (expressed positive as work done). As might be expected, paraffin wax and cetyl palmitate have nearly the same value for w which increases considerably for polar solid surfaces both for water and mercury.

When $\theta_R(\text{min.}) = 0^\circ$, i.e. where the liquid moves over the solid leaving a film behind, w is indeterminate.

In table 4 (last column) w has been expressed (as H) in gcal./gmol. where possible to enable comparison with the heats of adsorption of similar types of systems from vapour solid studies. In the calculations a number of simplifying assumptions are made, e.g. the molecules are assumed to be perfect spheres, mercury to be monatomic and the density in the interfacial layer to be the same as in the liquid. The assumptions are valid if we are only concerned with the order of magnitude. For water the radius of the area occupied by the molecule in the interfacial layer is thus 2·4 Å. This value is close to that of 2·2 Å calculated by Bernal and Fowler (1933).

H varies from about 145 to 538 gcal./gmol. for water and from 483 to 1134 for mercury. The values for sodium hydroxide solutions on paraffin wax are calculated assuming that the interfacial layer contains only water; this is largely true if sodium hydroxide is negatively adsorbed at the liquid/wax interface as it is at the liquid/air interface. On this assumption, as might be expected, the values for 0·2N and 1·0N sodium hydroxide solutions are about the same as for pure water on the same surface (paraffin wax). The results for water on paraffin wax and cetyl palmitate are nearly the same. This is expected since the surfaces of paraffin wax and cetyl palmitate will be similar, for the latter's ester group lies between two long hydrocarbon chains; similarly for mercury on paraffin wax and cetyl palmitate.

The values obtained are of the order of van der Waals' forces, the lower ones are associated with non-polar surfaces (paraffin wax, cetyl palmitate), giving only hydrocarbon chains at the surfaces; the higher values are associated with polar surfaces (stearic acid and glass). No results can be calculated for the linseed oil/solid systems for the oil is a mixture of unknown molecular dimensions.

All the results obtained show that the forces involved in the difference between θ_A(const.) and θ_R(min.) are those corresponding to physical adsorption.

The variable advancing angle and the constant advancing angle

Table 3 shows those surfaces for which the advancing angle is constant or variable. It appears, however (table 5), that for any liquid/solid system there is a characteristic value θ_A(const.) though this may only be attained *immediately before the drop moves* over the unwetted solid.

A scrutiny of the surfaces for which θ_A is constant and those for which θ_A changes with α reveals an important difference. Whereas the former are relatively non-polar surfaces (paraffin wax, cetyl palmitate*) or composed of close-knit polar molecules (glass), the latter are polar surfaces of long chain compounds with polar heads and long hydrocarbon tails (stearic acid, tristearin, cetyl alcohol) or compounds whose polar groups are separated by hydrocarbon type groupings (gum arabic, manucol). This suggests the following explanation for the cases where θ_A varies with α. Generally the polar heads (COOH, OH, ester, etc.) will be more readily wetted by water, mercury, linseed oil, etc., than the non-polar portions of the molecule. If there are relatively few polar groups jutting out of the solid surface, then a liquid drop placed on the surface may, as a result of the forceful action of laying down, 'splash' before taking up its normal configuration and thus, on spreading or contraction, adhere to a larger number of polar heads than is justified by the average distribution of the polar groups on the surface. Since the adhesion for the polar groups is stronger a non-equilibrium 'pseudo' advancing angle θ_A^P may result $< \theta_A$(const.), the advancing angle really representative of the average surface configuration. This follows since W_{SL} (for θ_A^P) $> W_{SL}$ (for θ_A const.). Now θ_A^P represents a thermodynamically more stable state than θ_A(const.), therefore the advancing angle will not change (i.e. increase) to the normal value of θ_A(const.) unless energy is put into the system. However, on tilting, the drop can take up gravitational potential energy. This 'energy input' cannot be simply equated to the different energy states represented by θ_A^P and θ_A(const.), for the former will be inconstant and depend upon the degree of 'splashing' when the drop is put on the plate. Table 6 shows that for any given system θ_A^P can have many values as might be expected from the above hypothesis. However these variations in the advancing angle do not invalidate the linear relation of $(\cos\theta_R - \cos\theta_A)$ and $A \sin\alpha$ predicted by equation (2), for this involves only the difference of the contact angles at any angle of tilt.

---

* As previously indicated cetyl palmitate is composed of two long hydrocarbon chains joined by an intermediate ester group and its surface properties therefore resemble paraffin wax.

TABLE 6

liquid/solid	variation of the initial advancing angle (θ_A^P) at $\alpha = 0°$
water/stearic acid	82°, 84·5°, 77·5°, 81°, 87°, 85°
mercury/cetyl alcohol	146·2°, 151·2°, 149·7°
mercury/stearic acid	144·5°, 146·5°

It is of interest (table 7) that the differences between the lowest value of θ_A^P observed and θ_A(const.) is greater for those surfaces whose polar groups are attached to long hydrocarbon tails (stearic acid, cetyl alcohol, tristearin) than it is for the surfaces whose polar groups are separated by short chains (manucol, gum arabic). This is to be expected on the 'splash' hypothesis for the latter will be more closely knit.

TABLE 7. COMPARISON BETWEEN LOWEST θ_A^P AND θ_A(const.)

liquid/solid	θ_A^P lowest	θ_A(const.)
water/stearic acid	77·5°	105°
water/tristearin	112°	(120°)*
mercury/stearic acid	144·5°	155°
mercury/tristearin	148°	160°
mercury/cetyl alcohol	146·2°	153·5°
linseed oil/gum arabic	24·7°	28°
linseed oil/manucol	35·7°	39°

* Run not completed.

Water on monofilm of ferric stearate

According to Blodgett and Langmuir (1937) a stable monofilm is produced by 'rubbing down' ferric stearate on glass. It is probable, from electron diffraction studies, that the molecules are fixed with their polar heads attached to the glass and the hydrocarbon tails stick out (Germer 1939; Germer and Storks 1939). If this is true, such a monofilm should behave like a non-polar surface (see table 1) and θ_A should be constant and not vary with α. Table 8 shows this is the case for three representative films. All the monofilms do not give the same value of θ_A(const.) and they also give a slightly low value for the surface tension of water ($69 \pm 1{\cdot}5$ ergs/cm.^{2}). This is probably because the surfaces may easily become slightly contaminated by the rather crude 'rubbing down' process to which they are subjected.

TABLE 8. WATER ON FERRIC STEARATE MONOFILM

I		II		III	
α	θ_A	α	θ_A	α	θ_A
0°	86·1°	0°	93·0°	0°	97·5°
1·3°	86·3°	1·6°	93·2°	1·2°	97·6°
3·0°	86·0°	3·3°	93·4°	2·6°	98·0°
4·9°	86·0°	5·2°	93·8°	5·1°	98·6°
6·9°	86·1°	6·8°	93·9°	6·8°	97·9°
8·1°	86·1°	8·1°	93·5°	8·7°	98·4°
9·6°	86·0°	10·0°	93·6°	10·9°	98·2°
12·2°	85·6°	12·3°	93·7°		

Contact angles on built-up layers of monofilms

Langmuir (1938) found that the advancing and receding angles of a water drop on monolayers or built-up monolayers differ on tilting. An *S* layer of oleic acid on glass or a *B* layer of stearic acid on chromium plated brass gave initial contact angles of 55°; on tilting θ_A rose to about 90° and θ_R was about 40°. For a water drop on a barium stearate *Y* film (hydrocarbon tails sticking out) θ initial was 90°, but on tilting the plate until the drop began to move and then arresting this movement by backward rotation, the contact angles became $\theta_A = 110°$ and $\theta_R = 82°$. From these, and other observations, Langmuir postulates that 'hysteresis' in the advancing and receding angles is due to overturning and re-orientation of the film molecules due to the attraction of their polar groups by water, and asserts that this is strong evidence for the existence of this overturning and re-orientation. Thus θ_A and θ_R differ because the advancing edge of the drop moves on to a surface having few hydrophilic groups, while the receding edge peels back from a surface containing many hydrophilic groups arising from overturning and re-orientation by the attraction of the water drop for the film molecules right underneath it.

This is clearly not a general explanation for the difference between θ_A and θ_R since this also occurs for non-polar solids like paraffin wax (no hydrophilic groups). Further, the phenomenon is partially reversible on rotating backwards again, i.e. θ_R (hindmost edge) increases again although the plate is still inclined to the horizontal. On Langmuir's hypothesis this should be because the drop is tending to move backwards on to a surface with fewer hydrophilic groups. However, since the drop is still inclined to the horizontal, the hypothesis can hardly be true, for this implies that the drop is capable of moving *upwards* over a surface containing fewer hydrophilic groups and therefore for which the adhesion is weaker. The essential distinction between Langmuir's explanation for the difference in θ_A and θ_R and that advanced here is that the former postulates a re-orientation of the polar molecules of the surface layer of the film (and therefore cannot apply to non-polar surfaces), whereas the latter suggests that the difference is due to the peeling off of an *adsorbed layer of the liquid drop* at the liquid/film boundary.

Table 8 shows, for water on ferric stearate monofilm, θ_A is independent of α, i.e. $\theta_A = \theta_A(\text{const.})$; while on Langmuir's films the initial θ_A corresponds to θ_A^P. This may be explained if the vigorous rubbing down process anchors the ferric stearate monofilm more firmly and less randomly to the slide than is the case with Langmuir's films.* If this is the case the 'splash' hypothesis accounts for the variations of θ_A with α on Langmuir's films.

Langmuir, Schaefer and Sobotka (1937) and Langmuir and Schaefer (1937) employ static contact angle methods to elucidate the surface structure of built-up monofilm layers. They have also investigated qualitatively the angles of tilt at which drops of liquid fall off these surfaces. More valuable data may be obtained about these surfaces if the *differences* between θ_A and θ_R of suitable liquids were determined in the sense described here.

Bikerman (1939) suggested that on a hydrophobic multifilm surface the following relation holds: $W \tan \alpha = K$, where W = weight of the drop of liquid on the surface and α is the angle of inclination of the plate at which the drop slides off. In fact only for large values of W is $1/W$ approximately linear against $\tan \alpha$. For water drops on paraffin wax, $W \tan \alpha$ is a minimum when W is approximately 0·00049 g. Still smaller drops adhere to a vertical surface so $W \tan \alpha$ can rise to infinity. Linearity therefore only applies over a limited range.

The authors are indebted to the Printing and Allied Trades Research Association for permission to publish this paper.

---

* Langmuir's B and Y layers are produced by dipping slides under monolayers on water, and the S films by evaporation of solvent from solutions on slides.

References

Ablett 1923 *Phil. Mag.* **46**, 244.
Adam 1938 *The Physics and Chemistry of Surfaces*, pp. 181 and 164. Oxford Univ. Press.
Adam and Jessop 1925 *J. Chem. Soc.* p. 1865.
Bartell and Wooley 1933 *J. Amer. Chem. Soc.* p. 3518.
Bernal and Fowler 1933 *J. Chem. Phys.* **1**, 515.
Bikerman 1939 *Proc. Roy. Soc.* A, **170**, 130.
Blodgett and Langmuir 1937 *Phys. Rev.* **51**, 964.
Dupré 1869 *Théorie Mécanique de la Chaleur*, p. 369.
Edser 1922 *4th Colloid Report, Brit. Ass.* p. 292.
Germer 1939 Recent Advances in Surface Chemistry and Chemical Physics. *Rep. Amer. Ass. Sci.* **7**, 47.
Germer and Storks 1939 *Phys. Rev.* **55**, 648.
Langmuir 1938 *Science*, **87**, 493.
Langmuir and Schaefer 1937 *J. Amer. Chem. Soc.* **59**, 2400.
Langmuir, Schaefer and Sobotka 1937 *J. Amer. Chem. Soc.* **59**, 1751.
Pockels 1914 *Phys. Z.* **15**, 39.
Sulman 1920 *Trans. Inst. Min. Metall.* **29**, 88.
Young 1805 *Phil. Trans.* **95**, 84.

Stress systems in aeolotropic plates. IV

By A. E. Green, Ph.D.

(*Communicated by G. I. Taylor, F.R.S.—Received* 31 *July* 1941)

Stress distributions in an aeolotropic plate containing a circular hole are discussed theoretically when the material of the plate has two directions of symmetry at right angles to one another. Some examples of stress distributions are included which have non-zero force resultants on the edge of the hole, corresponding to cases in isotropic materials for which the solution is dependent on Poisson's ratio. The use of the complex variable makes the method of solution comparatively simple, and as an introduction to the work for an aeolotropic material the same method is applied to problems of stresses in an isotropic plate containing a circular hole in order to obtain results which Bickley previously found by another method. Numerical work is carried out using the elastic constants found in experiments with specimens cut from the highly aeolotropic materials spruce and oak.

Introduction

1. The problem of any stress distribution in an infinite isotropic plate containing a circular hole has been solved in a general manner by Bickley (1928). The results are greatly modified when the plate is no longer isotropic, and in a previous paper (Green and Taylor 1941) the stress distribution in an infinite aeolotropic tension member containing a circular hole was found, the material of the plate having two directions of symmetry at right angles.

The present paper is an extension of this work. A general solution is obtained for problems of stress distributions in an infinite aeolotropic plate containing a circular hole. The particular cases which are considered numerically are the stress distributions due to a uniform shear and bending moment; and those due to a uniform normal pressure and to a uniform twist applied at the edge of the hole. Also the stress distributions due to a normal pressure on the hole proportional to $\cos\theta$ and a tangential stress on the hole proportional to $\sin\theta$. These last two problems are slightly more complicated than the others as they produce non-zero force resultants on the edge of the hole and therefore correspond to cases in isotropic materials for which the solution is not independent of Poisson's ratio. A combination of some of these solutions is taken to give some idea of the stress distribution round a rivet driven into a hole and pulled sideways. All the numerical work is carried out for certain specimens of oak and spruce wood.

The method of solution which was used in the tension problem (Green and Taylor 1941) can be extended and applied to the general case. This method, which is analogous to that used by Bickley (1928) for isotropic materials, consists in finding suitable fundamental stress functions which satisfy the equations of equilibrium and which give single-valued expressions for the corresponding stresses and displacements, and then combining these functions in an infinite series so as to satisfy the boundary conditions. It is found that the form of the solution which is obtained by this process is not always completely satisfactory owing to difficulties of convergence when the material is highly aeolotropic. In the tension problem the solution was modified and expressed in a finite form which was completely satisfactory, and this modification may be carried out for the general problem. It has been found, however, that by using the complex variable it is possible to obtain the required results by a more satisfactory direct method which can be used both for isotropic and aeolotropic materials. As an introduction to the main work of this paper, Bickley's results for an isotropic material are obtained in § 2 and then the general solution for an aeolotropic material is given in § 4.

SOLUTION FOR ISOTROPIC MATERIALS

2. It is known that the mean stresses can be expressed in terms of a stress-function χ by the equations

$$\widehat{xx} = \frac{\partial^2\chi}{\partial y^2}, \quad \widehat{yy} = \frac{\partial^2\chi}{\partial x^2}, \quad \widehat{xy} = -\frac{\partial^2\chi}{\partial x\,\partial y}, \tag{2·1}$$

or

$$\widehat{rr} = \frac{1}{r^2}\frac{\partial^2\chi}{\partial\theta^2}+\frac{1}{r}\frac{\partial\chi}{\partial r}, \quad \widehat{r\theta} = -\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial\chi}{\partial\theta}\right), \quad \widehat{\theta\theta} = \frac{\partial^2\chi}{\partial r^2}, \tag{2·2}$$

provided that

$$\nabla_1^4\chi = 0. \tag{2·3}$$

The formulae for the mean displacements are, in rectangular co-ordinates,

$$u = \frac{1}{E}\frac{\partial\psi}{\partial y} - \frac{1+\sigma}{E}\frac{\partial\chi}{\partial x}, \quad v = \frac{1}{E}\frac{\partial\psi}{\partial x} - \frac{1+\sigma}{E}\frac{\partial\chi}{\partial y}, \tag{2·4}$$

where ψ is to be found from the equations

$$\nabla_1^2\psi = 0, \quad \frac{\partial^2\psi}{\partial x\,\partial y} = \nabla_1^2\chi. \tag{2·5}$$

E is Young's modulus and σ is Poisson's ratio.

A general solution of equation (2·3) may be taken to be the *real* part of

$$\chi = f(z) + \bar{z}g(z), \tag{2·6}$$

where
$$z = x+iy, \quad \bar{z} = x-iy, \tag{2·7}$$

and $f(z)$ and $g(z)$ are general functions of z. Although the stresses are obtained from the real part of (2·6), the use of the complex stress function greatly simplifies the algebra. The stress system corresponding to this stress function is

$$\left.\begin{aligned} \widehat{xx} &= -f''(z) + 2g'(z) - \bar{z}g''(z), \\ \widehat{yy} &= f''(z) + 2g'(z) + \bar{z}g''(z), \\ \widehat{xy} &= -if''(z) - i\bar{z}g''(z), \\ \widehat{rr} &= -\frac{zf''(z)}{\bar{z}} + 2g'(z) - zg''(z), \\ \widehat{\theta\theta} &= \frac{zf''(z)}{\bar{z}} + 2g'(z) + zg''(z), \\ \widehat{r\theta} &= -\frac{izf''(z)}{\bar{z}} - izg''(z), \end{aligned}\right\} \tag{2·8}$$

where dashes denote differentiation with respect to z. Also

$$\frac{\partial\psi}{\partial x} = -4ig(z), \quad \frac{\partial\psi}{\partial y} = 4g(z), \tag{2·9}$$

and the corresponding displacements are

$$\left.\begin{aligned} Eu &= (3-\sigma)g(z) - (1+\sigma)\{f'(z) + \bar{z}g'(z)\}, \\ Ev &= -i(3-\sigma)g(z) - i(1+\sigma)\{f'(z) + \bar{z}g'(z)\}. \end{aligned}\right\} \tag{2·10}$$

It is now necessary to consider what restrictions must be placed on the functions $f(z)$ and $g(z)$. The stress function (2·6) must produce single-valued stresses which have prescribed normal and tangential values on the edge of the circular hole $|z| = a$ and which tend to zero at infinity, and the corresponding displacements must be single-valued. From (2·8) and (2·10) it can be seen that the conditions at infinity and the conditions that the stresses and displacements must be single-valued may be satisfied by taking the stress function to be the *real* part of

$$\chi = f(z) + \bar z g(z) - \frac{(3-\sigma)(A-iB)z\log z}{1+\sigma} + (A+iB)\bar z \log z, \qquad (2\cdot 11)$$

where A and B are real constants, and $f'(z)$ and $g(z)$ are regular functions of z outside the circle $|z| = a$ which tend to zero at infinity. The stresses corresponding to this modified stress-function are

$$\widehat{xx} = -f''(z) + 2g'(z) - \bar z g''(z) + \frac{(5+\sigma)A}{(1+\sigma)z} - \frac{(1-3\sigma)iB}{(1+\sigma)z} + \frac{(A+iB)\bar z}{z^2}, \qquad (2\cdot 12)$$

$$\widehat{yy} = f''(z) + 2g'(z) + \bar z g''(z) - \frac{(1-3\sigma)A}{(1+\sigma)z} + \frac{(5+\sigma)iB}{(1+\sigma)z} - \frac{(A+iB)\bar z}{z^2}, \qquad (2\cdot 13)$$

$$\widehat{xy} = -if''(z) - i\bar z g''(z) + \frac{(3-\sigma)i(A-iB)}{(1+\sigma)z} + \frac{i(A+iB)\bar z}{z^2}, \qquad (2\cdot 14)$$

$$\widehat{rr} = -\frac{zf''(z)}{\bar z} + 2g'(z) - zg''(z) + \frac{3(A+iB)}{z} + \frac{(3-\sigma)(A-iB)}{(1+\sigma)\bar z}, \qquad (2\cdot 15)$$

$$\widehat{\theta\theta} = \frac{zf''(z)}{\bar z} + 2g'(z) + zg''(z) + \frac{A+iB}{z} - \frac{(3-\sigma)(A-iB)}{(1+\sigma)\bar z}, \qquad (2\cdot 16)$$

$$\widehat{r\theta} = -\frac{izf''(z)}{z} - izg''(z) + \frac{i(A+iB)}{z} + \frac{(3-\sigma)i(A-iB)}{(1+\sigma)\bar z}. \qquad (2\cdot 17)$$

The functions $f(z)$ and $g(z)$ and the constants A and B must be determined so that the values of the stresses $\widehat{rr}$ and $\widehat{r\theta}$ at the boundary of the hole agree with those prescribed. The most general formulae for the edge tractions may be taken to be

$$\left.\begin{aligned} -\widehat{rr}_e &= R_0 + \sum_{n=1}^{\infty} (R_n \cos n\theta + S_n \sin n\theta),\\ -\widehat{r\theta}_e &= T_0 + \sum_{n=1}^{\infty} (T_n \cos n\theta + U_n \sin n\theta).\end{aligned}\right\} \qquad (2\cdot 18)$$

In order to satisfy the boundary conditions it is convenient to introduce two functions of the complex variable $z = re^{i\theta}$ which Bickley (1928) defined by the series

$$\left.\begin{aligned} V = \phi + i\phi' &= \sum_{n=0}^{\infty} (R_n + iS_n) a^n z^{-n},\\ W = \psi + i\psi' &= \sum_{n=0}^{\infty} (-U_n + iT_n) a^n z^{-n},\end{aligned}\right\} \qquad (2\cdot 19)$$

or by the equivalent integral formulae

$$\left.\begin{aligned} V &= \frac{1}{2\pi}\int_0^{2\pi} \frac{ze^{-i\gamma}+a}{ze^{-i\gamma}-a}(\widehat{rr}_e)\,d\gamma,\\ W &= \frac{i}{2\pi}\int_0^{2\pi} \frac{ze^{-i\gamma}+a}{ze^{-i\gamma}-a}(\widehat{r\theta}_e)\,d\gamma,\end{aligned}\right\} \qquad (2\cdot 20)$$

where γ is the angular co-ordinate of the point of the boundary at which the values of $\widehat{rr}$ and $\widehat{r\theta}$ are taken. The functions V and W are finite at infinity and are such that the real part of $V = -\widehat{rr}_e$ and the imaginary part of $W = -\widehat{r\theta}_e$ on the edge of the hole. From (2·15) and (2·17) it is seen that V and W may also be written in the form

$$\left.\begin{aligned} V &= \frac{z^2 f''(z)}{a^2} + zg''(z) - 2g'(z) - \frac{3(A+iB)}{z} - \frac{(3-\sigma)(A+iB)}{(1+\sigma)z}, \\ W &= -\frac{z^2 f''(z)}{a^2} - zg''(z) + \frac{A+iB}{z} - \frac{(3-\sigma)(A+iB)}{(1+\sigma)z}. \end{aligned}\right\} \tag{2·21}$$

Since
$$V + W = -2g'(z) - \frac{2(A+iB)}{z} - \frac{2(3-\sigma)(A+iB)}{(1+\sigma)z} \tag{2·22}$$

does not contain any term which is independent of z, the unknown constants S_0 and U_0 in (2·19) must be chosen to be $-T_0$ and R_0 respectively.

The force resultant on the edge of the hole arises only from the terms in the stresses containing A and B and the components of the force resultant are given by

$$X = \int_0^{2\pi} (-\widehat{rr}_e \cos\theta + \widehat{r\theta}_e \sin\theta)\, a\, d\theta = -\frac{8\pi A}{1+\sigma}, \tag{2·23}$$

$$Y = -\int_0^{2\pi} (\widehat{rr}_e \sin\theta + \widehat{r\theta}_e \cos\theta)\, a\, d\theta = -\frac{8\pi B}{1+\sigma}. \tag{2·24}$$

By using the relation $\dfrac{\partial}{\partial\theta} = iz\dfrac{d}{dz}$ and the equations (2·15)–(2·17) and (2·19)–(2·24) and rejecting the imaginary parts of the various functions, it may easily be verified that the stresses in plane polar co-ordinates take the forms

$$\widehat{rr} = -(\phi+\psi) + \frac{a^2}{r^2}\psi + \frac{1}{2}\left(1 - \frac{a^2}{r^2}\right)\frac{\partial}{\partial\theta}(\phi'+\psi') + \frac{3-\sigma}{4\pi a}\frac{a}{r}\left(1 - \frac{a^2}{r^2}\right)(X\cos\theta + Y\sin\theta), \tag{2·25}$$

$$\widehat{\theta\theta} = -(\phi+\psi) - \frac{a^2}{r^2}\psi - \frac{1}{2}\left(1 - \frac{a^2}{r^2}\right)\frac{\partial}{\partial\theta}(\phi'+\psi') + \frac{3-\sigma}{4\pi a}\frac{a}{r}\left(1 + \frac{a^2}{r^2}\right)(X\cos\theta + Y\sin\theta), \tag{2·26}$$

$$\widehat{r\theta} = -\frac{a^2}{r^2}\psi' + \frac{1}{2}\left(1 - \frac{a^2}{r^2}\right)\frac{\partial}{\partial\theta}(\phi+\psi) + \frac{3-\sigma}{4\pi a}\frac{a}{r}\left(1 - \frac{a^2}{r^2}\right)(X\sin\theta - Y\cos\theta). \tag{2·27}$$

These expressions for the stresses are the same as those given by Bickley (1928) except that, owing to the way the constant terms in V and W have been chosen here, the terms containing the average pressure P on the boundary are contained in V and W in the present work instead of being displayed separately as in Bickley's results.

The above analysis may also be used for problems in which the hole is free from stress and stresses are transmitted from infinity.

FUNDAMENTAL EQUATIONS FOR AN AEOLOTROPIC MATERIAL

3. Rectangular co-ordinates x, y are taken in the plane of the plate parallel to the directions of symmetry of the material and with the origin at the centre of the circular hole. The mean stresses are still given by equations (2·1) and (2·2) but χ now satisfies the equation

$$\left(\frac{\partial^2}{\partial x^2}+\alpha_1\frac{\partial^2}{\partial y^2}\right)\left(\frac{\partial^2}{\partial x^2}+\alpha_2\frac{\partial^2}{\partial y^2}\right)\chi = 0, \tag{3·1}$$

where

$$\alpha_1\alpha_2 = s_{11}/s_{22}, \quad \alpha_1+\alpha_2 = (s_{66}+2s_{12})/s_{22}, \tag{3·2}$$

the elastic constants s_{11}, s_{22}, s_{12}, s_{66} being defined in the usual way (Green and Taylor 1939).* In addition it is convenient to write

$$\gamma_1 = \frac{\alpha_1^{\frac{1}{2}}-1}{\alpha_1^{\frac{1}{2}}+1}, \quad \gamma_2 = \frac{\alpha_2^{\frac{1}{2}}-1}{\alpha_2^{\frac{1}{2}}+1}, \tag{3·3}$$

and the analysis is confined to cases where α_1 and α_2 are real and positive so that γ_1 and γ_2 lie between -1 and 1.

The mean displacements are given by

$$\bar{u} = (s_{12}-s_{11})\frac{\partial\chi}{\partial x}+s_{11}\frac{\partial\psi}{\partial y}, \quad \bar{v} = (s_{12}-s_{22})\frac{\partial\chi}{\partial y}+s_{22}\frac{\partial\psi}{\partial x}, \tag{3·4}$$

where ψ is to be found from the equations

$$\frac{\partial^2\psi}{\partial x\partial y} = \nabla_1^2\chi, \quad \frac{\partial^2\psi}{\partial x^2}+\alpha_1\alpha_2\frac{\partial^2\psi}{\partial y^2} = (1-\alpha_1)(1-\alpha_2)\frac{\partial^2\chi}{\partial x\partial y}. \tag{3·5}$$

THE GENERAL METHOD OF SOLUTION

4. A general solution of equation (3·1) may be taken to be the *real* part of

$$\chi = f(z+\gamma_1\bar{z})+g(z+\gamma_2\bar{z}), \tag{4·1}$$

where z and $\bar{z}$ are defined in (2·7). The corresponding stress system is

$$\left.\begin{aligned}
\widehat{xx} &= -(1-\gamma_1)^2 f''(z+\gamma_1\bar{z})-(1-\gamma_2)^2 g''(z+\gamma_2\bar{z}),\\
\widehat{yy} &= \quad (1+\gamma_1)^2 f''(z+\gamma_1\bar{z})+(1+\gamma_2)^2 g''(z+\gamma_2\bar{z}),\\
\widehat{xy} &= -i(1-\gamma_1^2) f''(z+\gamma_1\bar{z})-i(1-\gamma_2^2) g''(z+\gamma_2\bar{z}),\\
\widehat{rr} &= -\frac{(z-\gamma_1\bar{z})^2}{z\bar{z}} f''(z+\gamma_1\bar{z})-\frac{(z-\gamma_2\bar{z})^2}{z\bar{z}} g''(z+\gamma_2\bar{z}),\\
\widehat{\theta\theta} &= \quad \frac{(z+\gamma_1\bar{z})^2}{z\bar{z}} f''(z+\gamma_1\bar{z})+\frac{(z+\gamma_2\bar{z})^2}{z\bar{z}} g''(z+\gamma_2\bar{z}),\\
\widehat{r\theta} &= -\frac{i(z^2-\gamma_1^2\bar{z}^2)}{z\bar{z}} f''(z+\gamma_1\bar{z})-\frac{i(z^2-\gamma_2^2\bar{z}^2)}{z\bar{z}} g''(z+\gamma_2\bar{z}).
\end{aligned}\right\} \tag{4·2}$$

---

* In Green (1939) the first formula of group (2) should read $e_{xx} = A\widehat{xx}+C\widehat{yy}$.

Also

$$\left.\begin{aligned}\frac{\partial\psi}{\partial x} &= -\frac{4i\gamma_1}{1-\gamma_1}f'(z+\gamma_1\bar{z})-\frac{4i\gamma_2}{1-\gamma_2}g'(z+\gamma_2\bar{z}),\\ \frac{\partial\psi}{\partial y} &= \frac{4\gamma_1}{1+\gamma_1}f'(z+\gamma_1\bar{z})+\frac{4\gamma_2}{1+\gamma_2}g'(z+\gamma_2\bar{z}),\end{aligned}\right\}\tag{4·3}$$

and the corresponding displacements are

$$\left.\begin{aligned}\bar{u} &= (1+\gamma_1)(s_{12}-\alpha_2 s_{22})f'(z+\gamma_1\bar{z})+(1+\gamma_2)(s_{12}-\alpha_1 s_{22})g'(z+\gamma_2\bar{z}),\\ \bar{v} &= i(1-\gamma_1)(s_{12}-\alpha_1 s_{22})f'(z+\gamma_1\bar{z})+i(1-\gamma_2)(s_{12}-\alpha_2 s_{22})g'(z+\gamma_2\bar{z}).\end{aligned}\right\}\tag{4·4}$$

The stress-function (4·1) must produce single-valued stresses which have prescribed normal and tangential values on the edge of the circular hole $|z| = a$ and which tend to zero at infinity, and the corresponding displacements must be single-valued. From (4·2) and (4·4) it can be seen that the conditions at infinity and the conditions that the stresses and displacements must be single-valued and finite may be satisfied by taking the stress function to be the real part of

$$\begin{aligned}\chi = f(z+\gamma_1\bar{z})&+\{A(1-\gamma_2)(s_{12}-\alpha_2 s_{22})+iB(1+\gamma_2)(s_{12}-\alpha_1 s_{22})\}\\ &\qquad\times(z+\gamma_1\bar{z})\log(z+\gamma_1\bar{z})\\ &+g(z+\gamma_2\bar{z})-\{A(1-\gamma_1)(s_{12}-\alpha_1 s_{22})+iB(1+\gamma_1)(s_{12}-\alpha_2 s_{22})\}\\ &\qquad\times(z+\gamma_2\bar{z})\log(z+\gamma_2\bar{z}),\end{aligned}\tag{4·5}$$

where A and B are real constants and $f'(z+\gamma_1\bar{z})$ and $g'(z+\gamma_2\bar{z})$ are regular functions of $z+\gamma_1\bar{z}$ and $z+\gamma_2\bar{z}$ respectively, outside the circle $|z| = a$, which are at least of $O\left(\frac{1}{z+\gamma_1\bar{z}}\right)$ and of $O\left(\frac{1}{z+\gamma_2\bar{z}}\right)$ respectively at infinity. The stresses corresponding to this new form of stress function are

$$\begin{aligned}\widehat{xx} = -(1-\gamma_1)^2\{f''(z+\gamma_1\bar{z})&+[A(1-\gamma_2)(s_{12}-\alpha_2 s_{22})\\ &+iB(1+\gamma_2)(s_{12}-\alpha_1 s_{22})]/(z+\gamma_1\bar{z})\}\\ -(1-\gamma_2)^2\{g''(z+\gamma_2\bar{z})&-[A(1-\gamma_1)(s_{12}-\alpha_1 s_{22})\\ &+iB(1+\gamma_1)(s_{12}-\alpha_2 s_{22})]/(z+\gamma_2\bar{z})\},\end{aligned}\tag{4·6}$$

$$\begin{aligned}\widehat{yy} = (1+\gamma_1)^2\{f''(z+\gamma_1\bar{z})&+[A(1-\gamma_2)(s_{12}-\alpha_2 s_{22})\\ &+iB(1+\gamma_2)(s_{12}-\alpha_1 s_{22})]/(z+\gamma_1\bar{z})\}\\ +(1+\gamma_2)^2\{g''(z+\gamma_2\bar{z})&-[A(1-\gamma_1)(s_{12}-\alpha_1 s_{22})\\ &+iB(1+\gamma_1)(s_{12}-\alpha_2 s_{22})]/(z+\gamma_2\bar{z})\},\end{aligned}\tag{4·7}$$

$$\begin{aligned}\widehat{xy} = -i(1-\gamma_1^2)\{f''(z+\gamma_1\bar{z})&+[A(1-\gamma_2)(s_{12}-\alpha_2 s_{22})\\ &+iB(1+\gamma_2)(s_{12}-\alpha_1 s_{22})]/(z+\gamma_1\bar{z})\}\\ -i(1-\gamma_2^2)\{g''(z+\gamma_2\bar{z})&-[A(1-\gamma_1)(s_{12}-\alpha_1 s_{22})\\ &+iB(1+\gamma_1)(s_{12}-\alpha_2 s_{22})]/(z+\gamma_2\bar{z})\},\end{aligned}\tag{4·8}$$

$$\widehat{rr} = -\frac{(z-\gamma_1\bar{z})^2}{z\bar{z}}\left\{f''(z+\gamma_1\bar{z})+\frac{A(1-\gamma_2)(s_{12}-\alpha_2 s_{22})+iB(1+\gamma_2)(s_{12}-\alpha_1 s_{22})}{z+\gamma_1\bar{z}}\right\}$$
$$-\frac{(z-\gamma_2\bar{z})^2}{z\bar{z}}\left\{g''(z+\gamma_2\bar{z})-\frac{A(1-\gamma_1)(s_{12}-\alpha_1 s_{22})+iB(1+\gamma_1)(s_{12}-\alpha_2 s_{22})}{z+\gamma_2\bar{z}}\right\}, \tag{4·9}$$

$$\widehat{\theta\theta} = \frac{(z+\gamma_1\bar{z})^2}{z\bar{z}}\left\{f''(z+\gamma_1\bar{z})+\frac{A(1-\gamma_2)(s_{12}-\alpha_2 s_{22})+iB(1+\gamma_2)(s_{12}-\alpha_1 s_{22})}{z+\gamma_1\bar{z}}\right\}$$
$$+\frac{(z+\gamma_2\bar{z})^2}{z\bar{z}}\left\{g''(z+\gamma_2\bar{z})-\frac{A(1-\gamma_1)(s_{12}-\alpha_1 s_{22})+iB(1+\gamma_1)(s_{12}-\alpha_2 s_{22})}{z+\gamma_2\bar{z}}\right\}, \tag{4·10}$$

$$\widehat{r\theta} = -\frac{i(z^2-\gamma_1^2\bar{z}^2)}{z\bar{z}}\left\{f''(z+\gamma_1\bar{z})+\frac{A(1-\gamma_2)(s_{12}-\alpha_2 s_{22})+iB(1+\gamma_2)(s_{12}-\alpha_1 s_{22})}{z+\gamma_1\bar{z}}\right\}$$
$$-\frac{i(z^2-\gamma_2^2\bar{z}^2)}{z\bar{z}}\left\{g''(z+\gamma_2\bar{z})-\frac{A(1-\gamma_1)(s_{12}-\alpha_1 s_{22})+iB(1+\gamma_1)(s_{12}-\alpha_2 s_{22})}{z+\gamma_2\bar{z}}\right\}. \tag{4·11}$$

The functions $f''(z+\gamma_1\bar{z})$ and $g''(z+\gamma_2\bar{z})$ and the constants A and B must now be determined so that the values of the stresses $\widehat{rr}$ and $\widehat{r\theta}$ at the boundary of the hole agree with those prescribed. The most general formulae for the edge tractions may be taken to be those given by equations (2·18). Functions V and W are again defined by equations (2·19) or (2·20) but equations (2·21) are now replaced by

$$V = \frac{1}{a^2}\left(z-\frac{\gamma_1 a^2}{z}\right)^2 f''\left(z+\frac{\gamma_1 a^2}{z}\right)+\frac{1}{a^2}\left(z-\frac{\gamma_2 a^2}{z}\right)^2 g''\left(z+\frac{\gamma_2 a^2}{z}\right)$$
$$+A(1-\gamma_2)(s_{12}-\alpha_2 s_{22})\left(\frac{1+\gamma_1}{z}-\frac{4\gamma_1}{z+\dfrac{\gamma_1 a^2}{z}}\right)$$
$$-iB(1+\gamma_2)(s_{12}-\alpha_1 s_{22})\left(\frac{1-\gamma_1}{z}+\frac{4\gamma_1}{z+\dfrac{\gamma_1 a^2}{z}}\right)$$
$$-A(1-\gamma_1)(s_{12}-\alpha_1 s_{22})\left(\frac{1+\gamma_2}{z}-\frac{4\gamma_2}{z+\dfrac{\gamma_2 a^2}{z}}\right)$$
$$+iB(1+\gamma_1)(s_{12}-\alpha_2 s_{22})\left(\frac{1-\gamma_2}{z}+\frac{4\gamma_2}{z+\dfrac{\gamma_2 a^2}{z}}\right), \tag{4·12}$$

$$W = -\frac{1}{a^2}\left(z^2-\frac{\gamma_1^2 a^4}{z^2}\right)f''\left(z+\frac{\gamma_1 a^2}{z}\right)-\frac{1}{a^2}\left(z^2-\frac{\gamma_2^2 a^4}{z^2}\right)g''\left(z+\frac{\gamma_2 a^2}{z}\right)+(A+iB)$$
$$\times\{(1+\gamma_1)(1-\gamma_2)(s_{12}-\alpha_2 s_{22})-(1-\gamma_1)(1+\gamma_2)(s_{12}-\alpha_1 s_{22})\}/z. \tag{4·13}$$

From (4·12) and (4·13) it may be seen that $V+W$ does not contain any term which is independent of z and so the constants S_0 and U_0 in (2·19) must still be chosen to be $-T_0$ and R_0 respectively.

Two new variables u and v are now defined by the equations

$$z = \tfrac{1}{2}u\left\{1+\left(1-\frac{4\gamma_1 a^2}{u^2}\right)^{\frac{1}{2}}\right\}, \quad z = \tfrac{1}{2}v\left\{1+\left(1-\frac{4\gamma_2 a^2}{v^2}\right)^{\frac{1}{2}}\right\}. \tag{4·14}$$

In order that z may be single-valued the square roots in these expressions are to be chosen by the following rule—the amplitudes ϕ_1 and ϕ_2 of $1-\dfrac{4\gamma_1 a^2}{u^2}$ and $1-\dfrac{4\gamma_2 a^2}{v^2}$ respectively are to lie between $\pm\pi$ and the amplitudes of the square roots of these quantities are to be $\tfrac{1}{2}\phi_1$ and $\tfrac{1}{2}\phi_2$ respectively.

Using (4·14), equations (4·12) and (4·13) give

$$\begin{aligned}
&(\gamma_1-\gamma_2)\left(1-\frac{4\gamma_1 a^2}{u^2}\right)^{\frac{1}{2}}\{uf''(u)+A(1-\gamma_2)(s_{12}-\alpha_2 s_{22})+iB(1+\gamma_2)(s_{12}-\alpha_1 s_{22})\}\\
&\quad = A(1-\gamma_2)^2(s_{12}-\alpha_2 s_{22})-A(1-\gamma_1)(1-\gamma_2)(s_{12}-\alpha_1 s_{22})\\
&\qquad -iB(1+\gamma_2)^2(s_{12}-\alpha_1 s_{22})+iB(1+\gamma_1)(1+\gamma_2)(s_{12}-\alpha_2 s_{22})\\
&\qquad -\frac{u}{4}\{V(u)+W(u)\}\left\{1+\left(1-\frac{4\gamma_1 a^2}{u^2}\right)^{\frac{1}{2}}\right\}\\
&\qquad -\frac{\gamma_2 u}{4\gamma_1}\{V(u)-W(u)\}\left\{1-\left(1-\frac{4\gamma_1 a^2}{u^2}\right)^{\frac{1}{2}}\right\},
\end{aligned} \tag{4·15}$$

$$\begin{aligned}
&(\gamma_1-\gamma_2)\left(1-\frac{4\gamma_2 a^2}{v^2}\right)^{\frac{1}{2}}\{vg''(v)-A(1-\gamma_1)(s_{12}-\alpha_1 s_{22})-iB(1+\gamma_1)(s_{12}-\alpha_2 s_{22})\}\\
&\quad = A(1-\gamma_1)^2(s_{12}-\alpha_1 s_{22})-A(1-\gamma_1)(1-\gamma_2)(s_{12}-\alpha_2 s_{22})\\
&\qquad -iB(1+\gamma_1)^2(s_{12}-\alpha_2 s_{22})+iB(1+\gamma_1)(1+\gamma_2)(s_{12}-\alpha_1 s_{22})\\
&\qquad +\frac{v}{4}\{V(v)+W(v)\}\left\{1+\left(1-\frac{4\gamma_2 a^2}{v^2}\right)^{\frac{1}{2}}\right\}\\
&\qquad +\frac{\gamma_1 v}{4\gamma_2}\{V(v)-W(v)\}\left\{1-\left(1-\frac{4\gamma_2 a^2}{v^2}\right)^{\frac{1}{2}}\right\}.
\end{aligned} \tag{4·16}$$

The force resultant on the edge of the hole arises only from terms in the stresses containing A and B, and the components of the force resultant are found to be

$$\left.\begin{aligned}
X &= 2\pi A(1-\gamma_1)(1-\gamma_2)(\alpha_1-\alpha_2)s_{22} = \pi a(R_1-U_1),\\
Y &= 2\pi B(1+\gamma_1)(1+\gamma_2)(\alpha_1-\alpha_2)s_{22} = \pi a(S_1+T_1).
\end{aligned}\right\} \tag{4·17}$$

For any particular problem the functions V and W are known and then the functions $f''(u)$ and $g''(v)$ may be calculated from (4·15) and (4·16), while the constants A and B are found from (4·17). The required stresses are then given by equations (4·6)–(4·11) where the variables u and v must now take the values

$$u = z+\gamma_1\bar{z}, \quad v = z+\gamma_2\bar{z}. \tag{4·18}$$

5. The above analysis may also be used for problems in which the hole is free from stress and the stresses take given values at infinity. If the hole is absent and the stresses are transmitted from infinity, there will be known values for these stresses on the circle $|z| = a$. A stress system which tends to zero at infinity and which cancels these stresses on the circle can be found by the method of § 4, and the complete stress system for the plate containing a hole free from stress is obtained by adding to this system the stresses transmitted from infinity.

6. In the rest of this paper particular examples of the above general theory will be considered. In previous papers on stress systems in aeolotropic materials numerical work was carried out for specimens of oak and spruce whose elastic constants are shown in table 1. These values are used for calculations in the present paper and they are chosen so that the grain of the wood is parallel to the y-axis. If the values of s_{11} and s_{22} are interchanged (which means a change in sign of γ_1 and γ_2), then the grain will be parallel to the x-axis. The wood is cut from the tree so that the annular layers are parallel to the plane of the board.

TABLE 1

	s_{11}	s_{22}	s_{12}	s_{66}	α_1	α_2	γ_1	γ_2
oak	10·15	1·72	−0·87	12·8	5·321	1·109	0·395	0·026
spruce	15·5	0·587	−0·33	11·5	16·91	1·56	0·608	0·111

For convenience the following notation is introduced, namely

$$X_1 = R_1^{\frac{1}{2}}e^{\frac{1}{2}i\phi_1}, \quad X_2 = R_2^{\frac{1}{2}}e^{\frac{1}{2}i\phi_2}, \tag{6·1}$$

where
$$R_1 e^{i\phi_1} = 1-\frac{4a^2\gamma_1}{(z+\gamma_1\bar{z})^2}, \quad R_2 e^{i\phi_2} = 1-\frac{4a^2\gamma_2}{(z+\gamma_2\bar{z})^2}. \tag{6·2}$$

The quantities R_1 and R_2 are chosen to be positive and ϕ_1 and ϕ_2 are to lie between $-\pi$ and π.

Although the tension problem has been fully discussed in a previous paper (Green and Taylor 1941), it is interesting to obtain the chief formulae for the stresses by the improved method of the present paper.

THE TENSION PROBLEM

7. A uniform tension T is applied to the plate at infinity parallel to the x-axis. If the hole is absent, the stresses in plane polar co-ordinates are given by

$$\widehat{rr} = \tfrac{1}{2}T(1+\cos 2\theta), \quad \widehat{r\theta} = -\tfrac{1}{2}T\sin 2\theta, \quad \widehat{\theta\theta} = \tfrac{1}{2}T(1-\cos 2\theta). \tag{7·1}$$

It is now necessary to find a stress system which tends to zero at infinity and which takes the values

$$-\widehat{rr}_e = \tfrac{1}{2}T(1+\cos 2\theta), \quad -\widehat{r\theta}_e = -\tfrac{1}{2}T\sin 2\theta, \tag{7·2}$$

on the edge of the hole $r = a$. Thus, from (2·19),

$$V = \tfrac{1}{2}T(1+a^2/z^2), \quad W = \tfrac{1}{2}T(-1+a^2/z^2), \tag{7·3}$$

and using (4·14) these become

$$V(u)-W(u) = V(v)-W(v) = T, \tag{7·4}$$

$$\left.\begin{aligned} V(u)+W(u) &= \frac{T\left\{1-\left(1-\dfrac{4\gamma_1 a^2}{u^2}\right)^{\frac{1}{2}}\right\}}{\gamma_1\left\{1+\left(1-\dfrac{4\gamma_1 a^2}{u^2}\right)^{\frac{1}{2}}\right\}},\\ V(v)+W(v) &= \frac{T\left\{1-\left(1-\dfrac{4\gamma_2 a^2}{v^2}\right)^{\frac{1}{2}}\right\}}{\gamma_2\left\{1+\left(1-\dfrac{4\gamma_2 a^2}{v^2}\right)^{\frac{1}{2}}\right\}}.\end{aligned}\right\} \tag{7·5}$$

Since the stresses (7·2) produce no resultant force on the hole, the constants A and B are zero and so, from (4·15), (4·16), (7·4) and (7·5), it is found that

$$\left.\begin{aligned} f''(u) &= \frac{T(1+\gamma_2)}{4\gamma_1(\gamma_1-\gamma_2)}\left\{1-\left(1-\frac{4\gamma_1 a^2}{u^2}\right)^{-\frac{1}{2}}\right\},\\ g''(v) &= -\frac{T(1+\gamma_1)}{4\gamma_2(\gamma_1-\gamma_2)}\left\{1-\left(1-\frac{4\gamma_2 a^2}{v^2}\right)^{-\frac{1}{2}}\right\},\end{aligned}\right\} \tag{7·6}$$

and the corresponding stresses are the real parts of the expressions (4·6)–(4·11) where u and v are now given by (4·18). Hence, adding in the stresses (7·1), the complete stress system is found from the real parts of the following functions:

$$\begin{aligned}\widehat{xx} = {} & \frac{(1+\gamma_1)(1+\gamma_2)T}{4\gamma_1\gamma_2}\\ & + \frac{T}{4(\gamma_1-\gamma_2)}\left\{\frac{(1+\gamma_2)(1-\gamma_1)^2}{\gamma_1 X_1} - \frac{(1+\gamma_1)(1-\gamma_2)^2}{\gamma_2 X_2}\right\},\end{aligned} \tag{7·7}$$

$$\begin{aligned}\widehat{yy} = {} & -\frac{(1+\gamma_1)(1+\gamma_2)T}{4\gamma_1\gamma_2}\\ & - \frac{T}{4(\gamma_1-\gamma_2)}\left\{\frac{(1+\gamma_2)(1+\gamma_1)^2}{\gamma_1 X_1} - \frac{(1+\gamma_1)(1+\gamma_2)^2}{\gamma_2 X_2}\right\},\end{aligned} \tag{7·8}$$

$$\widehat{xy} = \frac{iT}{4(\gamma_1-\gamma_2)}\left\{\frac{(1+\gamma_2)(1-\gamma_1^2)}{\gamma_1 X_1} - \frac{(1+\gamma_1)(1-\gamma_2^2)}{\gamma_2 X_2}\right\}, \tag{7·9}$$

$$\begin{aligned}\widehat{rr} = {} & \frac{(1+\gamma_1)(1+\gamma_2)}{4\gamma_1\gamma_2}T\cos 2\theta\\ & + \frac{T}{4(\gamma_1-\gamma_2)}\left\{\frac{(1+\gamma_2)(e^{i\theta}-\gamma_1 e^{-i\theta})^2}{\gamma_1 X_1} - \frac{(1+\gamma_1)(e^{i\theta}-\gamma_2 e^{-i\theta})^2}{\gamma_2 X_2}\right\},\end{aligned} \tag{7·10}$$

$$\widehat{\theta\theta} = -\frac{(1+\gamma_1)(1+\gamma_2)}{4\gamma_1\gamma_2} T\cos 2\theta$$
$$-\frac{T}{4(\gamma_1-\gamma_2)}\left\{\frac{(1+\gamma_2)(e^{i\theta}+\gamma_1 e^{-i\theta})^2}{\gamma_1 X_1} - \frac{(1+\gamma_1)(e^{i\theta}+\gamma_2 e^{-i\theta})^2}{\gamma_2 X_2}\right\}, \quad (7\cdot11)$$

$$\widehat{r\theta} = -\frac{(1+\gamma_1)(1+\gamma_2)}{4\gamma_1\gamma_2} T\sin 2\theta$$
$$+\frac{iT}{4(\gamma_1-\gamma_2)}\left\{\frac{(1+\gamma_2)(e^{2i\theta}-\gamma_1^2 e^{-2i\theta})}{\gamma_1 X_1} - \frac{(1+\gamma_1)(e^{2i\theta}-\gamma_2^2 e^{-2i\theta})}{\gamma_2 X_2}\right\}. \quad (7\cdot12)$$

The functions (7·7)–(7·12) are the complex conjugates of those given previously (Green and Taylor 1941) and so the values of the stresses found by the two methods agree since it is only the real parts of these functions which are required. The stresses at the edge of the hole $r = a$ take the simple form

$$\widehat{\theta\theta} = \frac{T(1+\gamma_1)(1+\gamma_2)(1+\gamma_1+\gamma_2-\gamma_1\gamma_2-2\cos 2\theta)}{(1+\gamma_1^2-2\gamma_1\cos 2\theta)(1+\gamma_2^2-2\gamma_2\cos 2\theta)}, \quad (7\cdot13)$$

$$\widehat{xx} = \widehat{\theta\theta}\sin^2\theta, \quad \widehat{yy} = \widehat{\theta\theta}\cos^2\theta, \quad \widehat{xy} = -\tfrac{1}{2}\widehat{\theta\theta}\sin 2\theta. \quad (7\cdot14)$$

THE SHEAR PROBLEM

8. In this problem the plate is acted on by uniform shears S at infinity parallel to the axes of symmetry. If the hole is absent, the stresses are

$$\widehat{rr} = S\sin 2\theta, \quad \widehat{r\theta} = S\cos 2\theta, \quad \widehat{\theta\theta} = -S\sin 2\theta, \quad (8\cdot1)$$

and so it is necessary to find a stress system which tends to zero at infinity and which takes the values

$$-\widehat{rr}_e = S\sin 2\theta, \quad -\widehat{r\theta}_e = S\cos 2\theta, \quad (8\cdot2)$$

on the edge of the hole $r = a$. Thus, from (2·19),

$$V = W = iSa^2/z^2, \quad (8\cdot3)$$

and using (4·14) these become

$$V(u)-W(u) = 0, \quad V(v)-W(v) = 0, \quad (8\cdot4)$$

$$\left.\begin{aligned} V(u)+W(u) &= \frac{2iS\left\{1-\left(1-\frac{4\gamma_1 a^2}{u^2}\right)^{\frac{1}{2}}\right\}}{\gamma_1\left\{1+\left(1-\frac{4\gamma_1 a^2}{u^2}\right)^{\frac{1}{2}}\right\}}, \\ V(v)+W(v) &= \frac{2iS\left\{1-\left(1-\frac{4\gamma_2 a^2}{v^2}\right)^{\frac{1}{2}}\right\}}{\gamma_2\left\{1+\left(1-\frac{4\gamma_2 a^2}{v^2}\right)^{\frac{1}{2}}\right\}}. \end{aligned}\right\} \quad (8\cdot5)$$

Since the stresses (8·2) produce no resultant force on the hole, the constants A and B are zero. Hence, from (4·15), (4·16), (8·4) and (8·5),

$$\left.\begin{aligned} f''(u) &= \frac{iS}{2\gamma_1(\gamma_1-\gamma_2)}\left\{1-\left(1-\frac{4\gamma_1 a^2}{u^2}\right)^{-\frac{1}{2}}\right\}, \\ g''(v) &= -\frac{iS}{2\gamma_2(\gamma_1-\gamma_2)}\left\{1-\left(1-\frac{4\gamma_2 a^2}{v^2}\right)^{-\frac{1}{2}}\right\}, \end{aligned}\right\} \tag{8·6}$$

and the corresponding stresses are the real parts of the expressions (4·6)–(4·11) where u and v are now given by (4·18). Therefore, adding in the stresses (8·1), the complete stress system is found from the real parts of

$$\widehat{xx} = \frac{iS}{2(\gamma_1-\gamma_2)}\left\{\frac{(1-\gamma_1)^2}{\gamma_1 X_1}-\frac{(1-\gamma_2)^2}{\gamma_2 X_2}\right\}, \tag{8·7}$$

$$\widehat{yy} = -\frac{iS}{2(\gamma_1-\gamma_2)}\left\{\frac{(1+\gamma_1)^2}{\gamma_1 X_1}-\frac{(1+\gamma_2)^2}{\gamma_2 X_2}\right\}, \tag{8·8}$$

$$\widehat{xy} = \frac{\gamma_1\gamma_2-1}{2\gamma_1\gamma_2}S-\frac{S}{2(\gamma_1-\gamma_2)}\left\{\frac{1-\gamma_1^2}{\gamma_1 X_1}-\frac{1-\gamma_2^2}{\gamma_2 X_2}\right\}, \tag{8·9}$$

$$\widehat{rr} = \frac{\gamma_1\gamma_2-1}{2\gamma_1\gamma_2}S\sin 2\theta+\frac{iS}{2(\gamma_1-\gamma_2)}\left\{\frac{(e^{i\theta}-\gamma_1 e^{-i\theta})^2}{\gamma_1 X_1}-\frac{(e^{i\theta}-\gamma_2 e^{-i\theta})^2}{\gamma_2 X_2}\right\}, \tag{8·10}$$

$$\widehat{\theta\theta} = -\frac{\gamma_1\gamma_2-1}{2\gamma_1\gamma_2}S\sin 2\theta-\frac{iS}{2(\gamma_1-\gamma_2)}\left\{\frac{(e^{i\theta}+\gamma_1 e^{-i\theta})^2}{\gamma_1 X_1}-\frac{(e^{i\theta}+\gamma_2 e^{-i\theta})^2}{\gamma_2 X_2}\right\}, \tag{8·11}$$

$$\widehat{r\theta} = \frac{\gamma_1\gamma_2-1}{2\gamma_1\gamma_2}S\cos 2\theta-\frac{S}{2(\gamma_1-\gamma_2)}\left\{\frac{e^{2i\theta}-\gamma_1^2 e^{-2i\theta}}{\gamma_1 X_1}-\frac{e^{2i\theta}-\gamma_2^2 e^{-2i\theta}}{\gamma_2 X_2}\right\}. \tag{8·12}$$

The stresses at the edge of the hole are of special interest and they take a comparatively simple form. Thus, when $r = a$,

$$\widehat{\theta\theta} = \frac{4S(\gamma_1\gamma_2-1)\sin 2\theta}{(1+\gamma_1^2-2\gamma_1\cos 2\theta)(1+\gamma_2^2-2\gamma_2\cos 2\theta)}, \tag{8·13}$$

$$\widehat{xx} = \widehat{\theta\theta}\sin^2\theta, \quad \widehat{yy} = \widehat{\theta\theta}\cos^2\theta, \quad \widehat{xy} = -\tfrac{1}{2}\widehat{\theta\theta}\sin 2\theta. \tag{8·14}$$

The results for an isotropic material can be deduced from equations (8·7)–(8·12) by considering their limit as $\gamma_1 \to \gamma_2 \to 0$. The well-known form for the stresses at the edge of a hole in an isotropic material are obtained immediately from (8·13) and (8·14) by putting $\gamma_1 = \gamma_2 = 0$.

The stresses at the edge of the hole have been evaluated numerically for the specimens of oak and of spruce whose elastic constants are given in table 1. The results are shown in table 2 for values of θ between 0° and 90° where the line $\theta = 0°$ is chosen to be parallel to the grain. The stresses at other points on the circle may easily be deduced from the given values. The corresponding numerical results for an isotropic material are included in table 2 for comparison.

Figures 1 and 2 illustrate the variations in $\widehat{\theta\theta}$ and $\widehat{xy}$ respectively over the edge of the hole for the given specimens of oak and spruce. The curves for an isotropic material are included for comparison. It will be seen that for

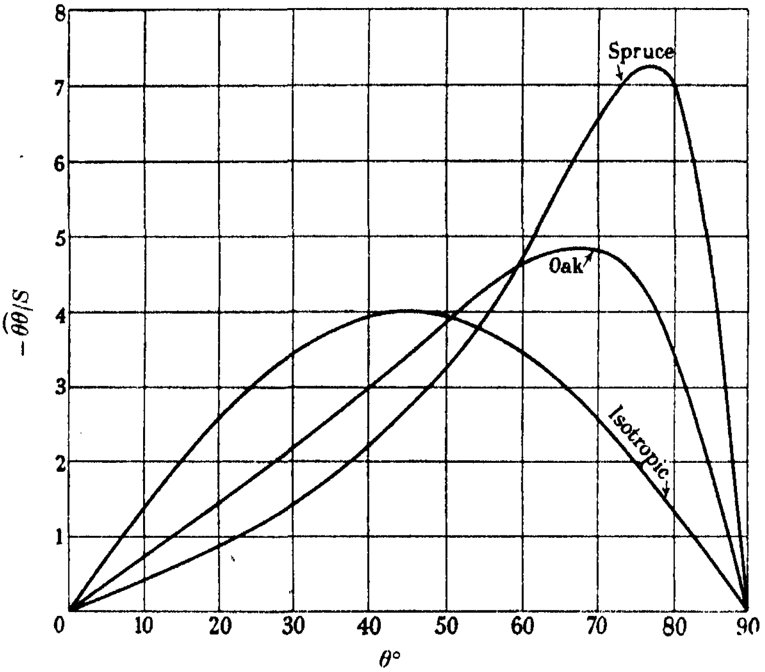


FIGURE 1. Distribution of $\widehat{\theta\theta}$ round the edge of a hole. Grain parallel to $\theta = 0$.

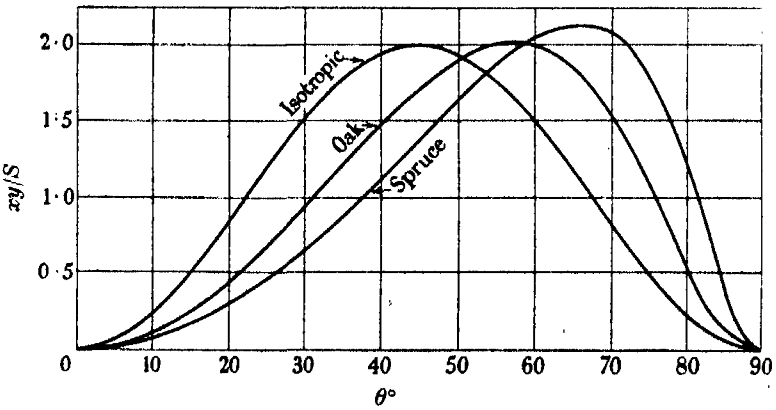


FIGURE 2. Distribution of $\widehat{xy}$ round the edge of a hole. Grain parallel to $\theta = 0$.

highly aeolotropic spruce the maximum circumferential stress rises to $\pm 7{\cdot}25S$ and that the positions of the maxima are at points lying about 13° on either side of the diameter of the hole which is perpendicular to the fibres. The corresponding maximum for an isotropic material is $\pm 4S$. The maximum shear stress however is approximately the same for spruce, oak and isotropic materials.

In the tension problem an attempt was made to apply the calculated stress concentrations in conjunction with measurements of ultimate strength to determine the kind of failure that might be expected to occur near a hole in a highly stressed spruce plate. A similar attempt for a plate under shear does not, however, lead to any significant results.

TABLE 2. SHEAR PROBLEM. VALUES OF STRESSES ON THE EDGE OF THE HOLE

θ°	$\widehat{xy}/S$	$-\widehat{xx}/S$	$-\widehat{yy}/S$	$-\widehat{\theta\theta}/S$
		Spruce		
0	0	0	0	0
10	0·0711	0·0125	0·403	0·416
20	0·283	0·103	0·777	0·881
30	0·630	0·364	1·09	1·45
40	1·09	0·914	1·30	2·21
50	1·61	1·91	1·35	3·26
60	2·04	3·53	1·18	4·71
70	2·09	5·74	0·760	6·50
75	1·80	6·71	0·482	7·20
80	1·20	6·79	0·211	7·00
85	0·412	4·70	0·036	4·74
90	0	0	0	0
		Oak		
0	0	0	0	0
10	0·116	0·0205	0·660	0·680
20	0·447	0·163	1·23	1·39
30	0·931	0·538	1·61	2·15
40	1·47	1·23	1·75	2·99
50	1·90	2·27	1·59	3·86
60	2·00	3·47	1·16	4·62
70	1·55	4·25	0·562	4·81
80	0·588	3·34	0·104	3·44
90	0	0	0	0
		Isotropic material		
0	0	0	0	0
10	0·234	0·0413	1·33	1·37
20	0·826	0·301	2·27	2·57
30	1·5	0·866	2·60	3·46
40	1·94	1·63	2·31	3·94
45	2·0	2·0	2·0	4·0

THE BENDING-MOMENT PROBLEM

9. In this section the plate is supposed to be bounded by two straight edges $y = \pm b$ where b is so large that the edges may be thought of as being at infinite distance apart. The plate is acted on by a bending moment so

that, if the hole is absent, there is a pure bending moment of amount M across every section of the plate parallel to the y-axis. In this case the stresses are

$$\left.\begin{aligned} \widehat{rr} = \frac{3Mr}{8b^3}(\sin\theta+\sin 3\theta), \quad \widehat{r\theta} &= -\frac{3Mr}{8b^3}(\cos\theta-\cos 3\theta),\\ \widehat{\theta\theta} &= \frac{3Mr}{8b^3}(3\sin\theta-\sin 3\theta), \end{aligned}\right\} \tag{9·1}$$

and it is now necessary to find a stress system which tends to zero at infinity and which takes the values

$$-\widehat{rr}_e = \frac{3Ma}{8b^3}(\sin\theta+\sin 3\theta), \quad -\widehat{r\theta}_e = -\frac{3Ma}{8b^3}(\cos\theta-\cos 3\theta), \tag{9·2}$$

at the edge of the hole $r = a$. Thus, from (2·19),

$$V = \frac{3Mai}{8b^3}\left(\frac{a}{z}+\frac{a^3}{z^3}\right), \quad W = \frac{3Mai}{8b^3}\left(-\frac{a}{z}+\frac{a^3}{z^3}\right), \tag{9·3}$$

and, using (4·14), these give

$$\left.\begin{aligned} V(u)-W(u) &= \frac{3Miu}{8b^3\gamma_1}\left\{1-\left(1-\frac{4\gamma_1a^2}{u^2}\right)^{\frac{1}{2}}\right\},\\ V(v)-W(v) &= \frac{3Miv}{8b^3\gamma_2}\left\{1-\left(1-\frac{4\gamma_2a^2}{v^2}\right)^{\frac{1}{2}}\right\}, \end{aligned}\right\} \tag{9·4}$$

$$\left.\begin{aligned} V(u)+W(u) &= \frac{3Miu\left\{1-\left(1-\frac{4\gamma_1a^2}{u^2}\right)^{\frac{1}{2}}\right\}^2}{8b^3\gamma_1^2\left\{1+\left(1-\frac{4\gamma_1a^2}{u^2}\right)^{\frac{1}{2}}\right\}},\\ V(v)+W(v) &= \frac{3Miv\left\{1-\left(1-\frac{4\gamma_2a^2}{v^2}\right)^{\frac{1}{2}}\right\}^2}{8b^3\gamma_2^2\left\{1+\left(1-\frac{4\gamma_2a^2}{v^2}\right)^{\frac{1}{2}}\right\}}. \end{aligned}\right\} \tag{9·5}$$

The constants A and B are zero and hence, from (4·15), (4·16), (9·4) and (9·5),

$$\left.\begin{aligned} f''(u) &= \frac{3Miu(1+\gamma_2)}{32b^3\gamma_1^2(\gamma_1-\gamma_2)}\left\{2-\left(1-\frac{4\gamma_1a^2}{u^2}\right)^{\frac{1}{2}}-\left(1-\frac{4\gamma_1a^2}{u^2}\right)^{-\frac{1}{2}}\right\},\\ g''(v) &= -\frac{3Miv(1+\gamma_1)}{32b^3\gamma_2^2(\gamma_1-\gamma_2)}\left\{2-\left(1-\frac{4\gamma_2a^2}{v^2}\right)^{\frac{1}{2}}-\left(1-\frac{4\gamma_2a^2}{v^2}\right)^{-\frac{1}{2}}\right\}, \end{aligned}\right\} \tag{9·6}$$

and the corresponding stresses are the real parts of the expressions (4·6)–(4·11) where u and v are now given by (4·18). Thus, when the stresses (9·1) are included, the complete system of stress is found from the real parts of the expressions

$$\begin{aligned}\frac{b^3\widehat{xx}}{M}=\frac{3(1+\gamma_1)(1+\gamma_2)(4\gamma_1\gamma_2-\gamma_1-\gamma_2)r\sin\theta}{16\gamma_1^2\gamma_2^2}\\+\frac{3ir}{32(\gamma_1-\gamma_2)}\{\gamma_1^{-2}(1+\gamma_2)(1-\gamma_1)^2(e^{i\theta}+\gamma_1e^{-i\theta})(X_1+X_1^{-1})\\-\gamma_2^{-2}(1+\gamma_1)(1-\gamma_2)^2(e^{i\theta}+\gamma_2e^{-i\theta})(X_2+X_2^{-1})\},\end{aligned}\tag{9·7}$$

$$\begin{aligned}\frac{b^3\widehat{yy}}{M}=\frac{3(1+\gamma_1)(1+\gamma_2)(\gamma_1+\gamma_2)r\sin\theta}{16\gamma_1^2\gamma_2^2}\\-\frac{3ir}{32(\gamma_1-\gamma_2)}\{\gamma_1^{-2}(1+\gamma_2)(1+\gamma_1)^2(e^{i\theta}+\gamma_1e^{-i\theta})(X_1+X_1^{-1})\\-\gamma_2^{-2}(1+\gamma_1)(1+\gamma_2)^2(e^{i\theta}+\gamma_2e^{-i\theta})(X_2+X_2^{-1})\},\end{aligned}\tag{9·8}$$

$$\begin{aligned}\frac{b^3\widehat{xy}}{M}=-\frac{3(1+\gamma_1)(1+\gamma_2)(\gamma_1+\gamma_2)r\cos\theta}{16\gamma_1^2\gamma_2^2}\\-\frac{3r}{32(\gamma_1-\gamma_2)}\{\gamma_1^{-2}(1+\gamma_2)(1-\gamma_1^2)(e^{i\theta}+\gamma_1e^{-i\theta})(X_1+X_1^{-1})\\-\gamma_2^{-2}(1+\gamma_1)(1-\gamma_2^2)(e^{i\theta}+\gamma_2e^{-i\theta})(X_2+X_2^{-1})\},\end{aligned}\tag{9·9}$$

$$\begin{aligned}\frac{b^3\widehat{rr}}{M}=\frac{3(1+\gamma_1)(1+\gamma_2)r}{16\gamma_1^2\gamma_2^2}\{\gamma_1\gamma_2\sin\theta+(\gamma_1\gamma_2-\gamma_1-\gamma_2)\sin3\theta\}\\+\frac{3ir}{32(\gamma_1-\gamma_2)}\{\gamma_1^{-2}(1+\gamma_2)(e^{i\theta}-\gamma_1e^{-i\theta})^2(e^{i\theta}+\gamma_1e^{-i\theta})(X_1+X_1^{-1})\\-\gamma_2^{-2}(1+\gamma_1)(e^{i\theta}-\gamma_2e^{-i\theta})^2(e^{i\theta}+\gamma_2e^{-i\theta})(X_2+X_2^{-1})\},\end{aligned}\tag{9·10}$$

$$\begin{aligned}\frac{b^3\widehat{\theta\theta}}{M}=\frac{3(1+\gamma_1)(1+\gamma_2)r}{16\gamma_1^2\gamma_2^2}\{3\gamma_1\gamma_2\sin\theta-(\gamma_1\gamma_2-\gamma_1-\gamma_2)\sin3\theta\}\\-\frac{3ir}{32(\gamma_1-\gamma_2)}\{\gamma_1^{-2}(1+\gamma_2)(e^{i\theta}+\gamma_1e^{-i\theta})^3(X_1+X_1^{-1})\\-\gamma_2^{-2}(1+\gamma_1)(e^{i\theta}+\gamma_2e^{-i\theta})^3(X_2+X_2^{-1})\},\end{aligned}\tag{9·11}$$

$$\begin{aligned}\frac{b^3\widehat{r\theta}}{M}=\frac{3(1+\gamma_1)(1+\gamma_2)r}{16\gamma_1^2\gamma_2^2}\{-\gamma_1\gamma_2\cos\theta+(\gamma_1\gamma_2-\gamma_1-\gamma_2)\cos3\theta\}\\-\frac{3r}{32(\gamma_1-\gamma_2)}\{\gamma_1^{-2}(1+\gamma_2)(e^{i\theta}+\gamma_1e^{-i\theta})^2(e^{i\theta}-\gamma_1e^{-i\theta})(X_1+X_1^{-1})\\-\gamma_2^{-2}(1+\gamma_1)(e^{i\theta}+\gamma_2e^{-i\theta})^2(e^{i\theta}-\gamma_2e^{-i\theta})(X_2+X_2^{-1})\}.\end{aligned}\tag{9·12}$$

The values of the stresses on the edge of the hole take the simple forms

$$\frac{b^3\widehat{\theta\theta}}{aM}=\frac{3(1+\gamma_1)(1+\gamma_2)\{(1+\gamma_1+\gamma_2)\sin\theta-\sin3\theta\}}{2(1+\gamma_1^2-2\gamma_1\cos2\theta)(1+\gamma_2^2-2\gamma_2\cos2\theta)},\tag{9·13}$$

$$\widehat{xx}=\widehat{\theta\theta}\sin^2\theta,\quad \widehat{yy}=\widehat{\theta\theta}\cos^2\theta,\quad \widehat{xy}=-\tfrac{1}{2}\widehat{\theta\theta}\sin2\theta.\tag{9·14}$$

The well-known results for an isotropic material may be obtained from the above formulae by making $\gamma_1 \to \gamma_2 \to 0$.

Numerical values for the stresses at the edge of the hole are given in table 3*a* for spruce and table 3*b* for oak. The distribution of $\widehat{\theta\theta}$ over one quadrant of the hole is shown in figure 3 for the case when there is a bending moment across sections which are perpendicular to the grain and in figure 4 for the case when there is a bending moment across sections which are parallel to the grain. The distribution of $\widehat{\theta\theta}$ for a hole in an isotropic sheet, namely

$$\frac{b^3\widehat{\theta\theta}}{aM} = \tfrac{3}{2}(\sin\theta - \sin 3\theta),$$

is also shown in figures 3 and 4 for comparison. In these figures the sheet is supposed to be acted on by a bending moment across sections which are perpendicular to $\theta = 0^\circ$ and $\widehat{\theta\theta}$ is positive where there is a tension and negative where the stress is compressive. It will be seen that with highly aeolotropic spruce the maximum stress rises to $\pm 5{\cdot}53aM/b^3$ when the bending-moment acts across sections which are perpendicular to the grain

TABLE 3*a*. VALUES OF STRESSES AT THE EDGE OF THE HOLE: SPRUCE

θ°	$\dfrac{b^3\widehat{xy}}{aM}$	$\dfrac{b^3\widehat{xx}}{aM}$	$\dfrac{b^3\widehat{yy}}{aM}$	$\dfrac{b^3\widehat{\theta\theta}}{aM}$
		Grain parallel to line $\theta = 0^\circ$		
0	0	0	0	0
10	0·0132	−0·00232	−0·0746	−0·0769
20	0·0475	−0·0173	−0·131	−0·148
30	0·0875	−0·0505	−0·152	−0·202
40	0·106	−0·0891	−0·127	−0·216
50	0·0649	−0·0774	−0·0545	−0·132
60	−0·0802	0·139	0·0463	0·185
70	−0·348	0·956	0·127	1·08
75	−0·493	1·84	0·132	1·97
80	−0·559	3·17	0·0986	3·27
85	−0·414	4·73	0·0362	4·76
90	0	5·53	0	5·53
		Grain perpendicular to line $\theta = 0^\circ$		
0	0	0	0	0
5	0·186	−0·0163	−2·12	−2·14
10	0·506	−0·0892	−2·87	−2·96
15	0·678	−0·182	−2·53	−2·71
20	0·649	−0·236	−1·78	−2·02
30	0·237	−0·137	−0·411	−0·548
40	−0·280	0·235	0·333	0·568
50	−0·649	0·774	0·545	1·32
60	−0·778	1·35	0·449	1·80
70	−0·669	1·84	0·244	2·08
80	−0·383	2·17	0·0674	2·24
90	0	2·28	0	2·28

TABLE 3*b*. VALUES OF STRESSES AT THE EDGE OF THE HOLE: OAK

$\theta°$	$\dfrac{b^3\widehat{xy}}{aM}$	$\dfrac{b^3\widehat{xx}}{aM}$	$\dfrac{b^3\widehat{yy}}{aM}$	$\dfrac{b^3\widehat{\theta\theta}}{aM}$
Grain parallel to line $\theta = 0°$				
0	0	0	0	0
10	0·0303	−0·00535	−0·172	−0·177
20	0·104	−0·0377	−0·285	−0·322
30	0·171	−0·0986	−0·296	−0·394
40	0·165	−0·138	−0·196	−0·334
50	0·0243	−0·0290	−0·0204	−0·0494
60	−0·259	0·448	0·149	0·597
70	−0·560	1·54	0·204	1·74
80	−0·551	3·12	0·0971	3·22
90	0	4·02	0	4·02
Grain perpendicular to line $\theta = 0°$				
0	0	0	0	0
10	0·236	−0·0416	−1·34	−1·38
20	0·495	−0·180	−1·36	−1·54
30	0·363	−0·209	−0·628	−0·838
40	−0·0495	0·0416	0·0590	0·101
50	−0·477	0·568	0·400	0·968
60	−0·718	1·24	0·415	1·66
70	−0·690	1·90	0·251	2·15
80	−0·417	2·37	0·0736	2·44
90	0	2·54	0	2·54

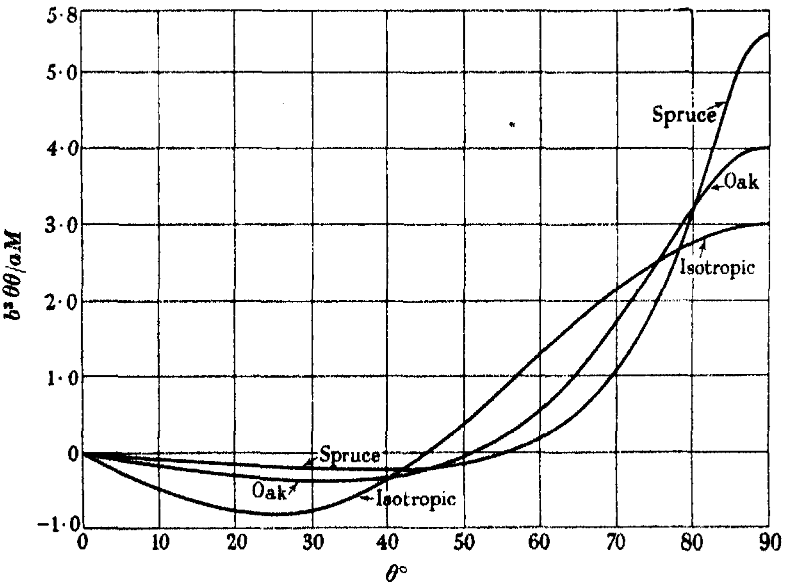


FIGURE 3. Stress distribution round the edge of a hole. Bending moment applied across sections which are perpendicular to the grain and to $\theta = 0$.

but that the region where the stress exceeds that for an isotropic material extends only 12° from the position of maximum stress concentrations at the points $\theta = \pm 90°$. When the bending moment acts across sections which are parallel to the grain, the maximum stress for spruce is $\pm 2{\cdot}96aM/b^3$ but the positions of the maxima are about 11° on either side of the line $\theta = 0°$.

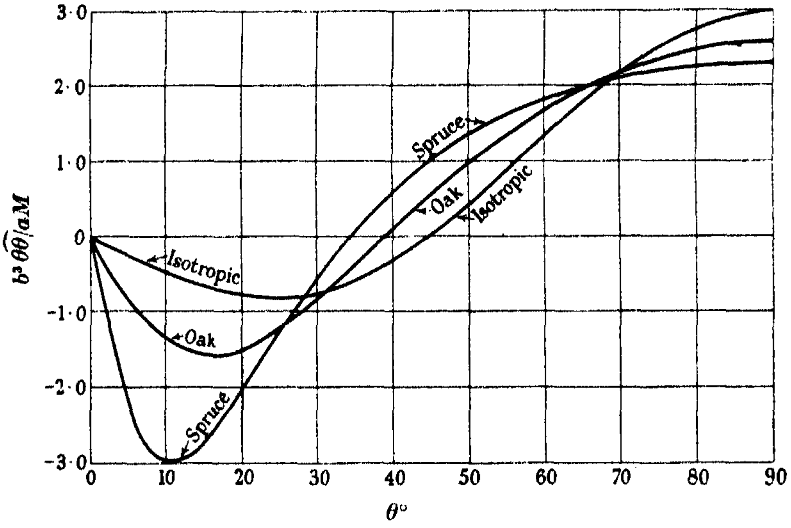


FIGURE 4. Stress distribution round the edge of a hole. Bending moment applied across sections which are parallel to the grain and perpendicular to $\theta = 0$.

HOLE UNDER ALL-ROUND PRESSURE

10. When the hole is acted on by a uniform normal pressure P, $R_0 = P$ and $V = P$, $W = -P$. The constants A and B are zero and

$$\left.\begin{aligned} f''(u) &= \frac{\gamma_2 P}{2\gamma_1(\gamma_1-\gamma_2)}\left\{1-\left(1-\frac{4\gamma_1 a^2}{u^2}\right)^{-\frac{1}{2}}\right\}, \\ g''(v) &= -\frac{\gamma_1 P}{2\gamma_2(\gamma_1-\gamma_2)}\left\{1-\left(1-\frac{4\gamma_2 a^2}{v^2}\right)^{-\frac{1}{2}}\right\}. \end{aligned}\right\} \qquad (10{\cdot}1)$$

The corresponding stresses are the real parts of the expressions (4·6)–(4·11) where u and v are given by (4·18). Thus, the stresses are the real parts of

$$\widehat{xx} = -P+\frac{\gamma_1+\gamma_2}{2\gamma_1\gamma_2}P+\frac{P}{2(\gamma_1-\gamma_2)}\left\{\frac{\gamma_2(1-\gamma_1)^2}{\gamma_1 X_1}-\frac{\gamma_1(1-\gamma_2)^2}{\gamma_2 X_2}\right\}, \qquad (10{\cdot}2)$$

$$\widehat{yy} = -P-\frac{\gamma_1+\gamma_2}{2\gamma_1\gamma_2}P-\frac{P}{2(\gamma_1-\gamma_2)}\left\{\frac{\gamma_2(1+\gamma_1)^2}{\gamma_1 X_1}-\frac{\gamma_1(1+\gamma_2)^2}{\gamma_2 X_2}\right\}, \qquad (10{\cdot}3)$$

$$\widehat{xy} = \frac{iP}{2(\gamma_1-\gamma_2)}\left\{\frac{\gamma_2(1-\gamma_1^2)}{\gamma_1 X_1}-\frac{\gamma_1(1-\gamma_2^2)}{\gamma_2 X_2}\right\}, \qquad (10{\cdot}4)$$

$$\widehat{rr} = -P + \frac{\gamma_1+\gamma_2}{2\gamma_1\gamma_2} P\cos 2\theta + \frac{P}{2(\gamma_1-\gamma_2)}\left\{\frac{\gamma_2(e^{i\theta}-\gamma_1 e^{-i\theta})^2}{\gamma_1 X_1} - \frac{\gamma_1(e^{i\theta}-\gamma_2 e^{-i\theta})^2}{\gamma_2 X_2}\right\}, \tag{10·5}$$

$$\widehat{\theta\theta} = -P - \frac{\gamma_1+\gamma_2}{2\gamma_1\gamma_2} P\cos 2\theta - \frac{P}{2(\gamma_1-\gamma_2)}\left\{\frac{\gamma_2(e^{i\theta}+\gamma_1 e^{-i\theta})^2}{\gamma_1 X_1} - \frac{\gamma_2(e^{i\theta}+\gamma_2 e^{-i\theta})^2}{\gamma_2 X_2}\right\}, \tag{10·6}$$

$$\widehat{r\theta} = -\frac{\gamma_1+\gamma_2}{2\gamma_1\gamma_2} P\sin 2\theta + \frac{iP}{2(\gamma_1-\gamma_2)}\left\{\frac{\gamma_2(e^{2i\theta}-\gamma_1^2 e^{-2i\theta})}{\gamma_1 X_1} - \frac{\gamma_1(e^{2i\theta}-\gamma_2^2 e^{-2i\theta})}{\gamma_2 X_2}\right\}. \tag{10·7}$$

When $r = a$ the stresses reduce to

$$\left.\begin{aligned} &\widehat{rr} = -P, \quad \widehat{r\theta} = 0, \\ &\widehat{\theta\theta} = P\frac{1+(\gamma_1+\gamma_2)^2-3\gamma_1^2\gamma_2^2-2(\gamma_1+\gamma_2)(1-\gamma_1\gamma_2)\cos 2\theta-2\gamma_1\gamma_2\cos 4\theta}{(1+\gamma_1^2-2\gamma_1\cos 2\theta)(1+\gamma_2^2-2\gamma_2\cos 2\theta)}, \end{aligned}\right\} \tag{10·8}$$

$$\left.\begin{aligned} &\widehat{xx} = -P\cos^2\theta+\widehat{\theta\theta}\sin^2\theta, \quad \widehat{yy} = -P\sin^2\theta+\widehat{\theta\theta}\cos^2\theta, \\ &\widehat{xy} = -\tfrac{1}{2}(P+\widehat{\theta\theta})\sin 2\theta. \end{aligned}\right\} \tag{10·9}$$

If $\gamma_1 \to \gamma_2 \to 0$, then $\widehat{rr} \to -Pa^2/r^2$, $\widehat{r\theta} \to 0$, $\widehat{\theta\theta} \to Pa^2/r^2$, which are the simple results for an isotropic material.

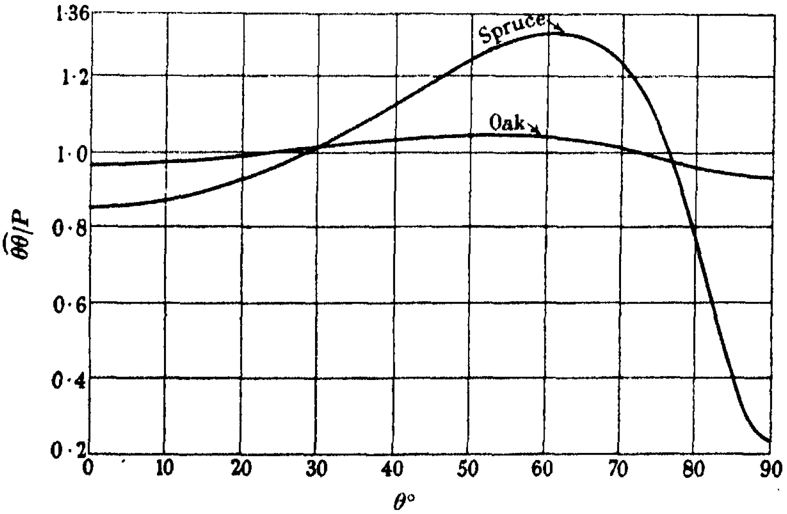


Figure 5. Stress distribution round the edge of a hole. Grain parallel to $\theta = 0$.

Numerical values for the stresses at the edge of the hole are given in table 4 and the distribution of $\widehat{\theta\theta}$ over one quadrant of the hole is shown in figure 5, where the line $\theta = 0°$ is parallel to the grain. It will be seen that the value of $\widehat{\theta\theta}/P$ for oak differs very little from the constant value 1 for an isotropic material. For spruce, however, $\widehat{\theta\theta}$ has a maximum of $1{\cdot}31P$ and a minimum of $0{\cdot}23P$, the minimum occurring at the two points on the circle where the normal pressure is perpendicular to the grain.

TABLE 4. VALUES OF STRESSES ON THE EDGE OF THE HOLE

$\theta°$	$\widehat{xy}/P$	$\widehat{xx}/P$	$\widehat{yy}/P$	$\widehat{\theta\theta}/P$
		Spruce		
0	0	−1·0	0·850	0·850
10	−0·319	−0·944	0·811	0·868
20	−0·618	−0·775	0·697	0·921
30	−0·870	−0·498	0·507	1·01
40	−1·05	−0·123	0·245	1·12
45	−1·09	0·090	0·090	1·18
50	−1·10	0·313	−0·0753	1·24
60	−1·00	0·733	−0·422	1·31
70	−0·716	0·968	−0·739	1·23
75	−0·516	0·924	−0·862	1·06
80	−0·303	0·715	−0·947	0·769
85	−0·122	0·399	−0·989	0·409
90	0	0·230	−1·0	0·230
		Oak		
0	0	−1·0	0·971	0·971
10	−0·338	−0·940	0·916	0·976
20	−0·639	−0·767	0·756	0·989
30	−0·869	−0·498	0·505	1·01
40	−0·998	−0·163	0·190	1·03
50	−1·01	0·198	−0·157	1·04
60	−0·883	0·529	−0·490	1·04
70	−0·646	0·776	−0·765	1·01
80	−0·335	0·901	−0·941	0·960
90	0	0·931	−1·0	0·931

ISOLATED TWIST AT THE HOLE

11. A uniform twist is applied at the edge of the hole so that it produces a couple G. Thus $T_0 = -\dfrac{G}{2\pi a^2}$, $V = \dfrac{iG}{2\pi a^2}$, $W = -\dfrac{iG}{2\pi a^2}$, the constants A and B are zero and

$$\left.\begin{aligned} f''(u) &= \frac{i\gamma_2 G}{4\pi\gamma_1 a^2(\gamma_1-\gamma_2)}\left\{1-\left(1-\frac{4\gamma_1 a^2}{u^2}\right)^{-\frac{1}{2}}\right\}, \\ g''(v) &= -\frac{i\gamma_1 G}{4\pi\gamma_2 a^2(\gamma_1-\gamma_2)}\left\{1-\left(1-\frac{4\gamma_2 a^2}{v^2}\right)^{-\frac{1}{2}}\right\}. \end{aligned}\right\} \tag{11·1}$$

The corresponding stresses are given by the expressions (4·6)–(4·11), where u and v are given by (4·18), and they are found to be the real parts of

$$\frac{2\pi a^2\widehat{xx}}{G} = \frac{i}{2(\gamma_1-\gamma_2)}\left\{\frac{\gamma_2(1-\gamma_1)^2}{\gamma_1 X_1} - \frac{\gamma_1(1-\gamma_2)^2}{\gamma_2 X_2}\right\}, \tag{11·2}$$

$$\frac{2\pi a^2\widehat{yy}}{G} = -\frac{i}{2(\gamma_1-\gamma_2)}\left\{\frac{\gamma_2(1+\gamma_1)^2}{\gamma_1 X_1} - \frac{\gamma_1(1+\gamma_2)^2}{\gamma_2 X_2}\right\}, \tag{11·3}$$

$$\frac{2\pi a^2\widehat{xy}}{G} = -\frac{\gamma_1+\gamma_2}{2\gamma_1\gamma_2} - \frac{1}{2(\gamma_1-\gamma_2)}\left\{\frac{\gamma_2(1-\gamma_1^2)}{\gamma_1 X_1} - \frac{\gamma_1(1-\gamma_2^2)}{\gamma_2 X_2}\right\}, \tag{11·4}$$

$$\frac{2\pi a^2\widehat{rr}}{G} = -\frac{\gamma_1+\gamma_2}{2\gamma_1\gamma_2}\sin 2\theta + \frac{i}{2(\gamma_1-\gamma_2)}\left\{\frac{\gamma_2(e^{i\theta}-\gamma_1 e^{-i\theta})^2}{\gamma_1 X_1} - \frac{\gamma_1(e^{i\theta}-\gamma_2 e^{-i\theta})^2}{\gamma_2 X_2}\right\}, \tag{11·5}$$

$$\frac{2\pi a^2\widehat{\theta\theta}}{G} = \frac{\gamma_1+\gamma_2}{2\gamma_1\gamma_2}\sin 2\theta - \frac{i}{2(\gamma_1-\gamma_2)}\left\{\frac{\gamma_2(e^{i\theta}+\gamma_1 e^{-i\theta})^2}{\gamma_1 X_1} - \frac{\gamma_1(e^{i\theta}+\gamma_2 e^{-i\theta})^2}{\gamma_2 X_2}\right\}, \tag{11·6}$$

$$\frac{2\pi a^2\widehat{r\theta}}{G} = -\frac{\gamma_1+\gamma_2}{2\gamma_1\gamma_2}\cos 2\theta - \frac{1}{2(\gamma_1-\gamma_2)}\left\{\frac{\gamma_2(e^{2i\theta}-\gamma_1^2 e^{-2i\theta})}{\gamma_1 X_1} - \frac{\gamma_1(e^{2i\theta}-\gamma_2^2 e^{-2i\theta})}{\gamma_2 X_2}\right\}. \tag{11·7}$$

The stresses at the edge of the hole are

$$2\pi a^2\widehat{r\theta} = G, \quad 2\pi a^2\widehat{\theta\theta} = \frac{4G\gamma_1\gamma_2\{(\gamma_1+\gamma_2)\sin 2\theta - \sin 4\theta\}}{(1+\gamma_1^2-2\gamma_1\cos 2\theta)(1+\gamma_2^2-2\gamma_2\cos 2\theta)}, \tag{11·8}$$

$$\left.\begin{aligned} \widehat{xx} &= \widehat{\theta\theta}\sin^2\theta - \widehat{r\theta}\sin 2\theta, \\ \widehat{yy} &= \widehat{\theta\theta}\cos^2\theta + \widehat{r\theta}\sin 2\theta, \\ \widehat{xy} &= -\tfrac{1}{2}\widehat{\theta\theta}\sin 2\theta + \widehat{r\theta}\cos 2\theta. \end{aligned}\right\} \tag{11·9}$$

When $\gamma_1 \to \gamma_2 \to 0$, $\widehat{rr} \to 0$, $\widehat{\theta\theta} \to 0$, $\widehat{r\theta} \to G/(2\pi r^2)$, which are the simple results for an isotropic material.

Numerical values for the stresses at the edge of a hole in specimens of oak and spruce are given in table 5 where the line $\theta = 0°$ is chosen to be parallel to the grain, and the distribution of $\dfrac{2\pi a^2\widehat{\theta\theta}}{G}$ over one quadrant of the hole is shown in figure 6. The value of $\dfrac{2\pi a^2\widehat{\theta\theta}}{G}$ for oak differs very little from zero which is the value for an isotropic material, but for spruce it attains a maximum value of about $\pm 0{\cdot}586$ at points which lie about 10° from the ends of the diameter of the hole which is perpendicular to the grain.

TABLE 5. VALUES OF STRESSES ON THE EDGE OF THE HOLE

$\theta°$	$\frac{2\pi a^2\widehat{xy}}{G}$	$\frac{2\pi a^2\widehat{xx}}{G}$	$\frac{2\pi a^2\widehat{yy}}{G}$	$\frac{2\pi a^2\widehat{\theta\theta}}{G}$
		Spruce		
0	1·0	0	0	0
10	0·953	−0·344	0·266	−0·0781
20	0·812	−0·660	0·516	−0·143
30	0·578	−0·911	0·731	−0·181
40	0·258	−1·06	0·885	−0·170
50	−0·131	−1·04	0·949	−0·0874
60	−0·541	−0·794	0·890	0·0954
70	−0·889	−0·306	0·687	0·382
75	−0·998	−0·00886	0·535	0·526
80	−1·04	0·226	0·360	0·586
85	−1·02	0·252	0·177	0·428
90	−1·0	0	0	0
		Oak		
0	1·0	0	0	0
10	0·943	−0·343	0·326	−0·0162
20	0·775	−0·646	0·618	−0·0281
30	0·514	−0·874	0·842	−0·0317
40	0·185	−0·995	0·971	−0·0238
50	−0·172	−0·987	0·984	−0·00295
60	−0·512	−0·845	0·873	0·0278
70	−0·784	−0·594	0·649	0·0554
80	−0·949	−0·292	0·344	0·0520
90	−1·0	0	0	0

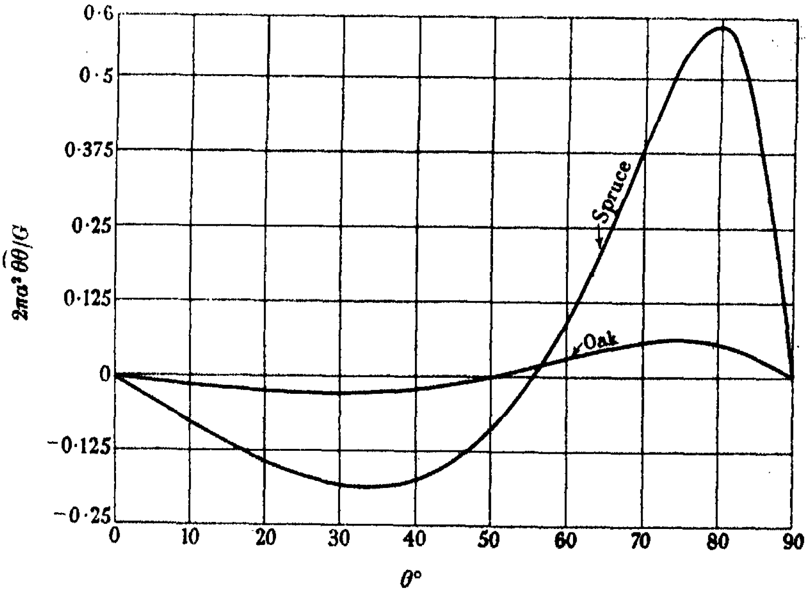


FIGURE 6. Stress distribution round the edge of a hole. Grain parallel to $\theta = 0$.

NORMAL PRESSURE VARYING AS COS θ AT THE EDGE OF THE HOLE

12. So far the examples of stress distributions which have been considered have been confined to those which form systems producing zero force resultants at the edge of the hole. The simplest case in which this condition is no longer true is when the hole is acted on by a normal pressure given by $\widehat{rr}_c = -P\cos\theta$ which produces a force resultant $X = \pi aP$ on the edge of the hole. In this case

$$2A(1-\gamma_1)(1-\gamma_2)(\alpha_1-\alpha_2)s_{22} = aP, \quad B = 0, \quad V = Pa/z, \quad W = 0, \tag{12·1}$$

so that, using (4·14),

$$V(u) = \frac{Pu}{2a\gamma_1}\left\{1-\left(1-\frac{4\gamma_1 a^2}{u^2}\right)^{\frac{1}{2}}\right\}, \quad V(v) = \frac{Pv}{2a\gamma_2}\left\{1-\left(1-\frac{4\gamma_2 a^2}{v^2}\right)^{\frac{1}{2}}\right\}. \tag{12·2}$$

Hence, from (4·15) and (4·16),

$$\begin{aligned} &f''(u) + A(1-\gamma_2)(s_{12}-\alpha_2 s_{22})/u \\ &\quad = \frac{aP(s_{12}-\alpha_2 s_{22})}{2(\alpha_1-\alpha_2)(1-\gamma_1)s_{22}}\,\frac{1}{u(1-4\gamma_1 a^2/u^2)^{\frac{1}{2}}} \\ &\qquad + \frac{\gamma_2 u}{8a\gamma_1^2(\gamma_1-\gamma_2)}\left\{2-\left(1-\frac{4\gamma_1 a^2}{u^2}\right)^{\frac{1}{2}}-\left(1-\frac{4\gamma_1 a^2}{u^2}\right)^{-\frac{1}{2}}\right\}, \end{aligned} \tag{12·3}$$

$$\begin{aligned} &g''(v) - A(1-\gamma_1)(s_{12}-\alpha_1 s_{22})/v \\ &\quad = -\frac{aP(s_{12}-\alpha_1 s_{22})}{2(\alpha_1-\alpha_2)(1-\gamma_2)s_{22}}\,\frac{1}{v(1-4\gamma_2 a^2/v^2)^{\frac{1}{2}}} \\ &\qquad - \frac{\gamma_1 v}{8a\gamma_2^2(\gamma_1-\gamma_2)}\left\{2-\left(1-\frac{4\gamma_2 a^2}{v^2}\right)^{\frac{1}{2}}-\left(1-\frac{4\gamma_2 a^2}{v^2}\right)^{-\frac{1}{2}}\right\}, \end{aligned} \tag{12·4}$$

and the corresponding stresses are the real parts of the expressions (4·6)–(4·11) where u and v are now given by (4·18). Thus, the stresses are given by the real parts of

$$\begin{aligned} \widehat{xx} = {} & \frac{\{\gamma_1^2+\gamma_2^2+\gamma_1\gamma_2-\gamma_1\gamma_2(\gamma_1+\gamma_2+\gamma_1\gamma_2)\}\,Pr\cos\theta}{4\gamma_1^2\gamma_2^2 a} \\ & - \frac{Pa}{2(\alpha_1-\alpha_2)s_{22}r}\left\{\frac{(s_{12}-\alpha_2 s_{22})(1-\gamma_1)}{(e^{i\theta}+\gamma_1 e^{-i\theta})X_1} - \frac{(s_{12}-\alpha_1 s_{22})(1-\gamma_2)}{(e^{i\theta}+\gamma_2 e^{-i\theta})X_2}\right\} \\ & + \frac{Pr}{8(\gamma_1-\gamma_2)a}\{\gamma_2\gamma_1^{-2}(1-\gamma_1)^2(e^{i\theta}+\gamma_1 e^{-i\theta})(X_1+X_1^{-1}) \\ & \qquad - \gamma_1\gamma_2^{-2}(1-\gamma_2)^2(e^{i\theta}+\gamma_2 e^{-i\theta})(X_2+X_2^{-1})\}, \end{aligned} \tag{12·5}$$

$$\begin{aligned} \widehat{yy} = {} & -\frac{\{\gamma_1^2+\gamma_2^2+\gamma_1\gamma_2+3\gamma_1\gamma_2(\gamma_1+\gamma_2+\gamma_1\gamma_2)\}\,Pr\cos\theta}{4\gamma_1^2\gamma_2^2 a} \\ & + \frac{Pa}{2(\alpha_1-\alpha_2)s_{22}r}\left\{\frac{(s_{12}-\alpha_2 s_{22})(1+\gamma_1)^2}{(1-\gamma_1)(e^{i\theta}+\gamma_1 e^{-i\theta})X_1} - \frac{(s_{12}-\alpha_1 s_{22})(1+\gamma_2)^2}{(1-\gamma_2)(e^{i\theta}+\gamma_2 e^{-i\theta})X_2}\right\} \\ & - \frac{Pr}{8(\gamma_1-\gamma_2)a}\{\gamma_2\gamma_1^{-2}(1+\gamma_1)^2(e^{i\theta}+\gamma_1 e^{-i\theta})(X_1+X_1^{-1}) \\ & \qquad - \gamma_1\gamma_2^{-2}(1+\gamma_2)^2(e^{i\theta}+\gamma_2 e^{-i\theta})(X_2+X_2^{-1})\}, \end{aligned} \tag{12·6}$$

$$
\begin{aligned}
\widehat{xy} = {} & \frac{\{\gamma_1\gamma_2(\gamma_1+\gamma_2+\gamma_1\gamma_2)-(\gamma_1^2+\gamma_2^2+\gamma_1\gamma_2)\}\,Pr\sin\theta}{4\gamma_1^2\gamma_2^2 a} \\
& -\frac{iPa}{2(\alpha_1-\alpha_2)\,s_{22}r}\left\{\frac{(s_{12}-\alpha_2 s_{22})\,(1+\gamma_1)}{(e^{i\theta}+\gamma_1 e^{-i\theta})\,X_1}-\frac{(s_{12}-\alpha_1 s_{22})\,(1+\gamma_2)}{(e^{i\theta}+\gamma_2 e^{-i\theta})\,X_2}\right\} \\
& +\frac{iPr}{8(\gamma_1-\gamma_2)\,a}\{\gamma_2\gamma_1^{-2}(1-\gamma_1^2)\,(e^{i\theta}+\gamma_1 e^{-i\theta})\,(X_1+X_1^{-1}) \\
& \qquad -\gamma_1\gamma_2^{-2}(1-\gamma_2^2)\,(e^{i\theta}+\gamma_2 e^{-i\theta})\,(X_2+X_2^{-1})\},
\end{aligned}
\tag{12·7}
$$

$$
\begin{aligned}
\widehat{rr} = {} & \frac{Pr}{4\gamma_1^2\gamma_2^2 a}\{-\gamma_1\gamma_2(\gamma_1+\gamma_2+\gamma_1\gamma_2)\cos\theta+(\gamma_1^2+\gamma_2^2+\gamma_1\gamma_2)\cos 3\theta\} \\
& -\frac{Pa}{2(\alpha_1-\alpha_2)\,s_{22}r}\left\{\frac{(s_{12}-\alpha_2 s_{22})\,(e^{i\theta}-\gamma_1 e^{-i\theta})^2}{(1-\gamma_1)\,(e^{i\theta}+\gamma_1 e^{-i\theta})\,X_1}-\frac{(s_{12}-\alpha_1 s_{22})\,(e^{i\theta}-\gamma_2 e^{-i\theta})^2}{(1-\gamma_2)\,(e^{i\theta}+\gamma_2 e^{-i\theta})\,X_2}\right\} \\
& +\frac{Pr}{8(\gamma_1-\gamma_2)\,a}\{\gamma_2\gamma_1^{-2}(e^{i\theta}-\gamma_1 e^{-i\theta})^2\,(e^{i\theta}+\gamma_1 e^{-i\theta})\,(X_1+X_1^{-1}) \\
& \qquad -\gamma_1\gamma_2^{-2}(e^{i\theta}-\gamma_2 e^{-i\theta})^2\,(e^{i\theta}+\gamma_2 e^{-i\theta})\,(X_2+X_2^{-1})\},
\end{aligned}
\tag{12·8}
$$

$$
\begin{aligned}
\widehat{\theta\theta} = {} & -\frac{Pr}{4\gamma_1^2\gamma_2^2 a}\{3\gamma_1\gamma_2(\gamma_1+\gamma_2+\gamma_1\gamma_2)\cos\theta+(\gamma_1^2+\gamma_2^2+\gamma_1\gamma_2)\cos 3\theta\} \\
& +\frac{Pa}{2(\alpha_1-\alpha_2)\,s_{22}r}\left\{\frac{(s_{12}-\alpha_2 s_{22})\,(e^{i\theta}+\gamma_1 e^{-i\theta})}{(1-\gamma_1)\,X_1}-\frac{(s_{12}-\alpha_1 s_{22})\,(e^{i\theta}+\gamma_2 e^{-i\theta})}{(1-\gamma_2)\,X_2}\right\} \\
& -\frac{Pr}{8(\gamma_1-\gamma_2)\,a}\{\gamma_2\gamma_1^{-2}(e^{i\theta}+\gamma_1 e^{-i\theta})^3\,(X_1+X_1^{-1}) \\
& \qquad -\gamma_1\gamma_2^{-2}(e^{i\theta}+\gamma_2 e^{-i\theta})^3\,(X_2+X_2^{-1})\},
\end{aligned}
\tag{12·9}
$$

$$
\begin{aligned}
\widehat{r\theta} = {} & -\frac{Pr}{4\gamma_1^2\gamma_2^2 a}\{\gamma_1\gamma_2(\gamma_1+\gamma_2+\gamma_1\gamma_2)\sin\theta+(\gamma_1^2+\gamma_2^2+\gamma_1\gamma_2)\sin 3\theta\} \\
& -\frac{iPa}{2(\alpha_1-\alpha_2)\,s_{22}r}\left\{\frac{(s_{12}-\alpha_2 s_{22})\,(e^{i\theta}-\gamma_1 e^{-i\theta})}{(1-\gamma_1)\,X_1}-\frac{(s_{12}-\alpha_1 s_{22})\,(e^{i\theta}-\gamma_2 e^{-i\theta})}{(1-\gamma_2)\,X_2}\right\} \\
& +\frac{iPr}{8(\gamma_1-\gamma_2)\,a}\{\gamma_2\gamma_1^{-2}(e^{i\theta}+\gamma_1 e^{-i\theta})^2\,(e^{i\theta}-\gamma_1 e^{-i\theta})\,(X_1+X_1^{-1}) \\
& \qquad -\gamma_1\gamma_2^{-2}(e^{i\theta}+\gamma_2 e^{-i\theta})^2(e^{i\theta}-\gamma_2 e^{-i\theta})\,(X_2+X_2^{-1})\}.
\end{aligned}
\tag{12·10}
$$

When $r = a$ the stresses are given by

$$
\begin{aligned}
\widehat{\theta\theta} = {} & -P\cos\theta \\
& +\frac{2+2\gamma_1\gamma_2-2\gamma_1\gamma_2(\gamma_1+\gamma_2)+2\gamma_1^2+2\gamma_2^2+4(\gamma_1\gamma_2-\gamma_1-\gamma_2)\cos 2\theta}{(1+\gamma_1^2-2\gamma_1\cos 2\theta)\,(1+\gamma_2^2-2\gamma_2\cos 2\theta)}\,P\cos\theta \\
& +\frac{2P\cos\theta}{(\alpha_1-\alpha_2)\,s_{22}}\left\{\frac{\gamma_1(s_{12}-\alpha_2 s_{22})}{1+\gamma_1^2-2\gamma_1\cos 2\theta}-\frac{\gamma_2(s_{12}-\alpha_1 s_{22})}{1+\gamma_2^2-2\gamma_2\cos 2\theta}\right\},
\end{aligned}
\tag{12·11}
$$

$$
\left.
\begin{aligned}
\widehat{xx} &= \tfrac{1}{2}(\widehat{rr}+\widehat{\theta\theta})+\tfrac{1}{2}(\widehat{rr}-\widehat{\theta\theta})\cos 2\theta, \\
\widehat{yy} &= \tfrac{1}{2}(\widehat{rr}+\widehat{\theta\theta})-\tfrac{1}{2}(\widehat{rr}-\widehat{\theta\theta})\cos 2\theta, \\
\widehat{xy} &= \tfrac{1}{2}(\widehat{rr}-\widehat{\theta\theta})\sin 2\theta.
\end{aligned}
\right\}
\tag{12·12}
$$

TABLE 6*a*. VALUES OF STRESSES AT THE EDGE OF THE HOLE: SPRUCE

$\theta°$	$\dfrac{\pi a\widehat{xy}}{X}$	$\dfrac{\pi a\widehat{xx}}{X}$	$\dfrac{\pi a\widehat{yy}}{X}$	$\dfrac{\pi a\widehat{\theta\theta}}{X}$
Grain parallel to resultant force				
0	0	−1·0	0·415	0·415
10	−0·238	−0·943	0·367	0·409
20	−0·426	−0·785	0·231	0·386
30	−0·519	−0·567	0·0323	0·332
40	−0·485	−0·359	−0·188	0·219
50	−0·319	−0·263	−0·375	0·00481
60	−0·0622	−0·3923	−0·464	−0·356
70	0·169	−0·807	−0·404	−0·869
75	0·214	−1·06	−0·316	−1·11
80	0·174	−1·16	−0·204	−1·19
85	0·0663	−0·845	−0·0930	−0·851
90	0	0	0	0
Grain perpendicular to resultant force				
0	0	−1·0	−0·161	−0·161
5	−0·0923	−0·988	0·0583	0·0664
10	−0·254	−0·940	0·454	0·499
15	−0·453	−0·844	0·726	0·848
20	−0·635	−0·709	0·804	1·035
30	−0·841	−0·381	0·591	1·08
40	−0·828	−0·0710	0·221	0·916
50	−0·663	0·148	−0·0863	0·704
60	−0·431	0·246	−0·251	0·494
70	−0·209	0·233	−0·266	0·309
80	−0·0550	0·138	−0·164	0·148
90	0	0	0	0

TABLE 6*b*. VALUES OF STRESSES AT THE EDGE OF THE HOLE: OAK

$\theta°$	$\dfrac{\pi a\widehat{xy}}{X}$	$\dfrac{\pi a\widehat{xx}}{X}$	$\dfrac{\pi a\widehat{yy}}{X}$	$\dfrac{\pi a\widehat{\theta\theta}}{X}$
Grain parallel to resultant force				
0	0	−1·0	0·442	0·442
10	−0·241	−0·942	0·384	0·427
20	−0·422	−0·786	0·220	0·374
30	−0·498	−0·579	−0·00381	0·284
40	−0·452	−0·387	−0·228	0·152
50	−0·307	−0·277	−0·385	−0·0192
60	−0·126	−0·282	−0·428	−0·210
70	0·00356	−0·352	−0·343	−0·353
80	0·0234	−0·306	−0·178	−0·310
90	0	0	0	0
Grain perpendicular to resultant force				
0	0	−1·0	0·212	0·212
10	−0·223	−0·946	0·281	0·320
20	−0·465	−0·771	0·337	0·506
30	−0·638	−0·498	0·240	0·608
40	−0·677	−0·198	0·0402	0·608
50	−0·581	0·0498	−0·155	0·537
60	−0·401	0·194	−0·269	0·425
70	−0·204	0·218	−0·268	0·292
80	−0·0550	0·138	−0·164	0·148
90	0	0	0	0

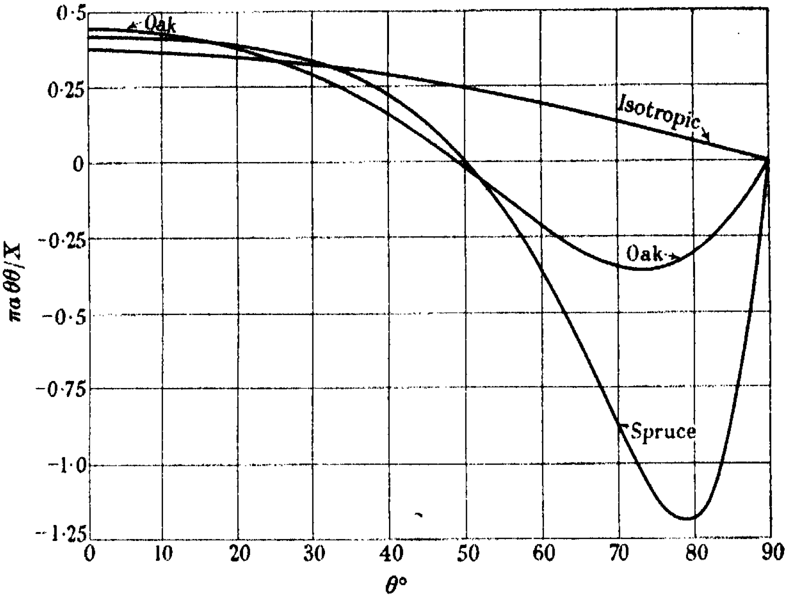


FIGURE 7. Stress distribution round the edge of a hole. Resultant force parallel to the grain and to $\theta = 0$.

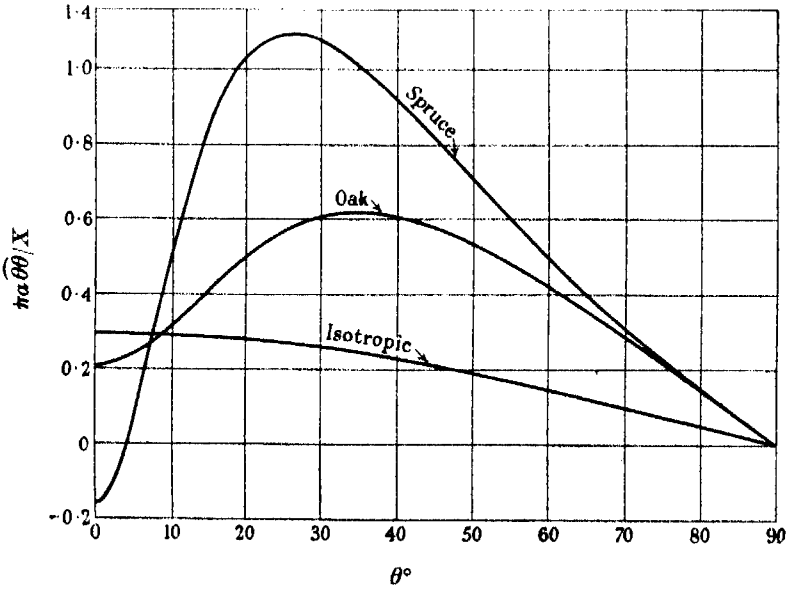


FIGURE 8. Stress distribution round the edge of a hole. Resultant force perpendicular to the grain and parallel to $\theta = 0$.

Numerical values for the stresses at the edge of the hole are given in table 6*a* for spruce and table 6*b* for oak. The distribution of $\widehat{\theta\theta}$ over one quadrant of the hole is shown in figure 7 for the case when the resultant force is parallel to the grain and in figure 8 when the resultant force is perpendicular to the grain. In each figure the resultant force is parallel to the line $\theta = 0$. The distribution of $\widehat{\theta\theta}$ for a hole in an isotropic sheet for which Poisson's ratio is 0·25 has been found by Bickley (1928) and is shown in figures 7 and 8 for comparison.

TANGENTIAL STRESS VARYING AS SIN θ AT THE EDGE OF THE HOLE

13. When a tangential stress $\widehat{r\theta}_e = P\sin\theta$ acts at the edge of the hole,

$$\left.\begin{aligned} &2A(1-\gamma_1)(1-\gamma_2)(\alpha_1-\alpha_2)s_{22} = aP, \\ &B = 0, \qquad V = 0, \qquad W = Pa/z, \end{aligned}\right\} \tag{13·1}$$

and, using (4·14),

$$W(u) = \frac{Pu}{2a\gamma_1}\left\{1-\left(1-\frac{4\gamma_1 a^2}{u^2}\right)^{\frac{1}{2}}\right\}, \quad W(v) = \frac{Pv}{2a\gamma_2}\left\{1-\left(1-\frac{4\gamma_2 a^2}{v^2}\right)^{\frac{1}{2}}\right\}. \tag{13·2}$$

Equations (4·15) and (4·16) then give

$$\begin{aligned} &f''(u) + A(1-\gamma_2)(s_{12}-\alpha_2 s_{22})/u \\ &\quad = \frac{aP(s_{12}-\alpha_2 s_{22})}{2(\alpha_1-\alpha_2)(1-\gamma_1)s_{22}}\,\frac{1}{u(1-4\gamma_1 a^2/u^2)^{\frac{1}{2}}} \\ &\qquad - \frac{\gamma_2 u}{8a\gamma_1^2(\gamma_1-\gamma_2)}\left\{2-\left(1-\frac{4\gamma_1 a^2}{u^2}\right)^{\frac{1}{2}}-\left(1-\frac{4\gamma_1 a^2}{u^2}\right)^{-\frac{1}{2}}\right\}, \end{aligned} \tag{13·3}$$

$$\begin{aligned} &g''(v) - A(1-\gamma_1)(s_{12}-\alpha_1 s_{22})/v \\ &\quad = -\frac{aP(s_{12}-\alpha_1 s_{22})}{2(\alpha_1-\alpha_2)(1-\gamma_2)s_{22}}\,\frac{1}{v(1-4\gamma_2 a^2/v^2)^{\frac{1}{2}}} \\ &\qquad + \frac{\gamma_1 v}{8a\gamma_2^2(\gamma_1-\gamma_2)}\left\{2-\left(1-\frac{4\gamma_2 a^2}{v^2}\right)^{\frac{1}{2}}-\left(1-\frac{4\gamma_2 a^2}{v^2}\right)^{-\frac{1}{2}}\right\}, \end{aligned} \tag{13·4}$$

and the corresponding stresses are the real parts of expressions (4·6)–(4·11) where u and v are now given by (4·18). Hence, the stresses are obtained from the real parts of

$$\begin{aligned} \widehat{xx} = {} & \frac{\{\gamma_1\gamma_2(\gamma_1+\gamma_2+\gamma_1\gamma_2)-(\gamma_1^2+\gamma_2^2+\gamma_1\gamma_2)\}\,Pr\cos\theta}{4\gamma_1^2\gamma_2^2 a} \\ & - \frac{Pa}{2(\alpha_1-\alpha_2)s_{22}r}\left\{\frac{(s_{12}-\alpha_2 s_{22})(1-\gamma_1)}{(e^{i\theta}+\gamma_1 e^{-i\theta})X_1} - \frac{(s_{12}-\alpha_1 s_{22})(1-\gamma_2)}{(e^{i\theta}+\gamma_2 e^{-i\theta})X_2}\right\} \\ & - \frac{Pr}{8(\gamma_1-\gamma_2)a}\{\gamma_2\gamma_1^{-2}(1-\gamma_1)^2(e^{i\theta}+\gamma_1 e^{-i\theta})(X_1+X_1^{-1}) \\ & \qquad\qquad - \gamma_1\gamma_2^{-2}(1-\gamma_2)^2(e^{i\theta}+\gamma_2 e^{-i\theta})(X_2+X_2^{-1})\}, \end{aligned} \tag{13·5}$$

$$\widehat{yy} = \frac{\{\gamma_1^2+\gamma_2^2+\gamma_1\gamma_2+3\gamma_1\gamma_2(\gamma_1+\gamma_2+\gamma_1\gamma_2)\}\,Pr\cos\theta}{4\gamma_1^2\gamma_2^2 a}$$

$$+\frac{Pa}{2(\alpha_1-\alpha_2)\,s_{22}r}\left\{\frac{(s_{12}-\alpha_2 s_{22})\,(1+\gamma_1)^2}{(1-\gamma_1)\,(e^{i\theta}+\gamma_1 e^{-i\theta})\,X_1}-\frac{(s_{12}-\alpha_1 s_{22})\,(1+\gamma_2)^2}{(1-\gamma_2)\,(e^{i\theta}+\gamma_2 e^{-i\theta})\,X_2}\right\}$$

$$+\frac{Pr}{8(\gamma_1-\gamma_2)\,a}\{\gamma_2\gamma_1^{-2}(1+\gamma_1)^2\,(e^{i\theta}+\gamma_1 e^{-i\theta})\,(X_1+X_1^{-1})$$

$$-\gamma_1\gamma_2^{-2}(1+\gamma_2)^2\,(e^{i\theta}+\gamma_2 e^{-i\theta})\,(X_2+X_2^{-1})\}, \tag{13·6}$$

$$\widehat{xy} = \frac{\{\gamma_1^2+\gamma_2^2+\gamma_1\gamma_2-\gamma_1\gamma_2(\gamma_1+\gamma_2+\gamma_1\gamma_2)\}\,Pr\sin\theta}{4\gamma_1^2\gamma_2^2 a}$$

$$-\frac{iPa}{2(\alpha_1-\alpha_2)\,s_{22}r}\left\{\frac{(s_{12}-\alpha_2 s_{22})\,(1+\gamma_1)}{(e^{i\theta}+\gamma_1 e^{-i\theta})\,X_1}-\frac{(s_{12}-\alpha_1 s_{22})\,(1+\gamma_2)}{(e^{i\theta}+\gamma_2 e^{-i\theta})\,X_2}\right\}$$

$$-\frac{iPr}{8(\gamma_1-\gamma_2)\,a}\{\gamma_2\gamma_1^{-2}(1-\gamma_1^2)\,(e^{i\theta}+\gamma_1 e^{-i\theta})\,(X_1+X_1^{-1})$$

$$-\gamma_1\gamma_2^{-2}(1-\gamma_2^2)\,(e^{i\theta}+\gamma_2 e^{-i\theta})\,(X_2+X_2^{-1})\}, \tag{13·7}$$

$$\widehat{rr} = \frac{Pr}{4\gamma_1^2\gamma_2^2 a}\{\gamma_1\gamma_2(\gamma_1+\gamma_2+\gamma_1\gamma_2)\cos\theta-(\gamma_1^2+\gamma_2^2+\gamma_1\gamma_2)\cos 3\theta\}$$

$$-\frac{Pa}{2(\alpha_1-\alpha_2)\,s_{22}r}\left\{\frac{(s_{12}-\alpha_2 s_{22})\,(e^{i\theta}-\gamma_1 e^{-i\theta})^2}{(1-\gamma_1)\,(e^{i\theta}+\gamma_1 e^{-i\theta})\,X_1}-\frac{(s_{12}-\alpha_1 s_{22})\,(e^{i\theta}-\gamma_2 e^{-i\theta})^2}{(1-\gamma_2)\,(e^{i\theta}+\gamma_2 e^{-i\theta})\,X_2}\right\}$$

$$-\frac{Pr}{8(\gamma_1-\gamma_2)\,a}\{\gamma_2\gamma_1^{-2}(e^{i\theta}-\gamma_1 e^{-i\theta})^2\,(e^{i\theta}+\gamma_1 e^{-i\theta})\,(X_1+X_1^{-1})$$

$$-\gamma_1\gamma_2^{-2}(e^{i\theta}-\gamma_2 e^{-i\theta})^2\,(e^{i\theta}+\gamma_2 e^{-i\theta})\,(X_2+X_2^{-1})\}, \tag{13·8}$$

$$\widehat{\theta\theta} = \frac{Pr}{4\gamma_1^2\gamma_2^2 a}\{3\gamma_1\gamma_2(\gamma_1+\gamma_2+\gamma_1\gamma_2)\cos\theta+(\gamma_1^2+\gamma_2^2+\gamma_1\gamma_2)\cos 3\theta\}$$

$$+\frac{Pa}{2(\alpha_1-\alpha_2)\,s_{22}r}\left\{\frac{(s_{12}-\alpha_2 s_{22})\,(e^{i\theta}+\gamma_1 e^{-i\theta})}{(1-\gamma_1)\,X_1}-\frac{(s_{12}-\alpha_1 s_{22})\,(e^{i\theta}+\gamma_2 e^{-i\theta})}{(1-\gamma_2)\,X_2}\right\}$$

$$+\frac{Pr}{8(\gamma_1-\gamma_2)\,a}\{\gamma_2\gamma_1^{-2}(e^{i\theta}+\gamma_1 e^{-i\theta})^3\,(X_1+X_1^{-1})$$

$$-\gamma_1\gamma_2^{-2}(e^{i\theta}+\gamma_2 e^{-i\theta})^3\,(X_2+X_2^{-1})\}, \tag{13·9}$$

$$\widehat{r\theta} = \frac{Pr}{4\gamma_1^2\gamma_2^2 a}\{\gamma_1\gamma_2(\gamma_1+\gamma_2+\gamma_1\gamma_2)\sin\theta+(\gamma_1^2+\gamma_2^2+\gamma_1\gamma_2)\sin 3\theta\}$$

$$-\frac{iPa}{2(\alpha_1-\alpha_2)\,s_{22}r}\left\{\frac{(s_{12}-\alpha_2 s_{22})\,(e^{i\theta}-\gamma_1 e^{-i\theta})}{(1-\gamma_1)\,X_1}-\frac{(s_{12}-\alpha_1 s_{22})\,(e^{i\theta}-\gamma_2 e^{-i\theta})}{(1-\gamma_2)\,X_2}\right\}$$

$$-\frac{iPr}{8(\gamma_1-\gamma_2)\,a}\{\gamma_2\gamma_1^{-2}(e^{i\theta}+\gamma_1 e^{-i\theta})^2\,(e^{i\theta}-\gamma_1 e^{-i\theta})\,(X_1+X_1^{-1})$$

$$-\gamma_1\gamma_2^{-2}(e^{i\theta}+\gamma_2 e^{-i\theta})^2\,(e^{i\theta}-\gamma_2 e^{-i\theta})\,(X_2+X_2^{-1})\}. \tag{13·10}$$

The stresses at the edge of the hole take the forms

$$\widehat{\theta\theta} = 2P\cos\theta$$

$$-\frac{2+2\gamma_1\gamma_2-2\gamma_1\gamma_2(\gamma_1+\gamma_2)+2\gamma_1^2+2\gamma_2^2+4(\gamma_1\gamma_2-\gamma_1-\gamma_2)\cos 2\theta}{(1+\gamma_1^2-2\gamma_1\cos 2\theta)\,(1+\gamma_2^2-2\gamma_2\cos 2\theta)}P\cos\theta$$

$$+\frac{2P\cos\theta}{(\alpha_1-\alpha_2)\,s_{22}}\left\{\frac{\gamma_1(s_{12}-\alpha_2 s_{22})}{1+\gamma_1^2-2\gamma_1\cos 2\theta}-\frac{\gamma_2(s_{12}-\alpha_1 s_{22})}{1+\gamma_2^2-2\gamma_2\cos 2\theta}\right\}, \tag{13·11}$$

$$\left.\begin{aligned} \widehat{xx} &= \widehat{\theta\theta}\sin^2\theta-\widehat{r\theta}\sin 2\theta,\\ \widehat{yy} &= \widehat{\theta\theta}\cos^2\theta+\widehat{r\theta}\sin 2\theta,\\ \widehat{xy} &= -\tfrac{1}{2}\widehat{\theta\theta}\sin 2\theta+\widehat{r\theta}\cos 2\theta.\end{aligned}\right\} \tag{13·12}$$

Numerical values for the stresses at the edge of the hole are given in table 7*a* for spruce and table 7*b* for oak. The distribution of $\widehat{\theta\theta}$ over one quadrant of the hole is shown in figure 9 for the case when the resultant force is parallel to the grain and in figure 10 for the case when the resultant

TABLE 7*a*. VALUES OF STRESSES AT THE EDGE OF THE HOLE: SPRUCE

$\theta°$	$\dfrac{\pi a\widehat{xy}}{X}$	$\dfrac{\pi a\widehat{xx}}{X}$	$\dfrac{\pi a\widehat{yy}}{X}$	$\dfrac{\pi a\widehat{\theta\theta}}{X}$
		Grain parallel to resultant force		
0	0	0	−0·437	−0·437
10	0·242	−0·0733	−0·388	−0·461
20	0·432	−0·282	−0·247	−0·529
30	0·524	−0·591	−0·0411	−0·632
40	0·481	−0·943	0·192	−0·751
50	0·290	−1·26	0·400	−0·859
60	−0·0311	−1·45	0·518	−0·928
70	−0·421	−1·43	0·495	−0·930
75	−0·617	−1·30	0·424	−0·878
80	−0·798	−1·06	0·314	−0·745
85	−0·941	−0·629	0·170	−0·459
90	−1·0	0	0	0
		Grain perpendicular to resultant force		
0	0	0	−0·401	−0·401
5	0·119	−0·0181	−0·369	−0·387
10	0·224	−0·0702	−0·287	−0·357
15	0·303	−0·151	−0·166	−0·317
20	0·344	−0·250	−0·0047	−0·254
30	0·297	−0·460	0·352	−0·108
40	0·0997	−0·623	0·647	0·0243
50	−0·188	−0·690	0·800	0·111
60	−0·496	−0·641	0·786	0·145
70	−0·761	−0·491	0·619	0·129
80	−0·938	−0·265	0·339	0·0741
90	−1·0	0	0	0

TABLE 7b. VALUES OF STRESSES AT THE EDGE OF THE HOLE: OAK

$\theta°$	$\dfrac{\pi a\widehat{xy}}{X}$	$\dfrac{\pi a\widehat{xx}}{X}$	$\dfrac{\pi a\widehat{yy}}{X}$	$\dfrac{\pi a\widehat{\theta\theta}}{X}$
		Grain parallel to resultant force		
0	0	0	−0·528	−0·528
10	0·255	−0·0756	−0·463	−0·539
20	0·444	−0·286	−0·279	−0·565
30	0·512	−0·584	−0·0202	−0·604
40	0·432	−0·902	0·252	−0·650
50	0·207	−1·16	0·469	−0·691
60	−0·127	−1·28	0·574	−0·706
70	−0·512	−1·18	0·528	−0·647
80	−0·853	−0·750	0·324	−0·426
90	−1·0	0	0	0
		Grain perpendicular to resultant force		
0	0	0	−0·716	−0·716
10	0·271	−0·0783	−0·549	−0·628
20	0·410	−0·274	−0·185	−0·459
30	0·382	−0·509	0·205	−0·304
40	0·203	−0·710	0·524	−0·186
50	−0·0817	−0·816	0·711	−0·104
60	−0·411	−0·788	0·737	−0·0504
70	−0·713	−0·622	0·602	−0·0201
80	−0·924	−0·343	0·337	−0·00581
90	−1·0	0	0	0

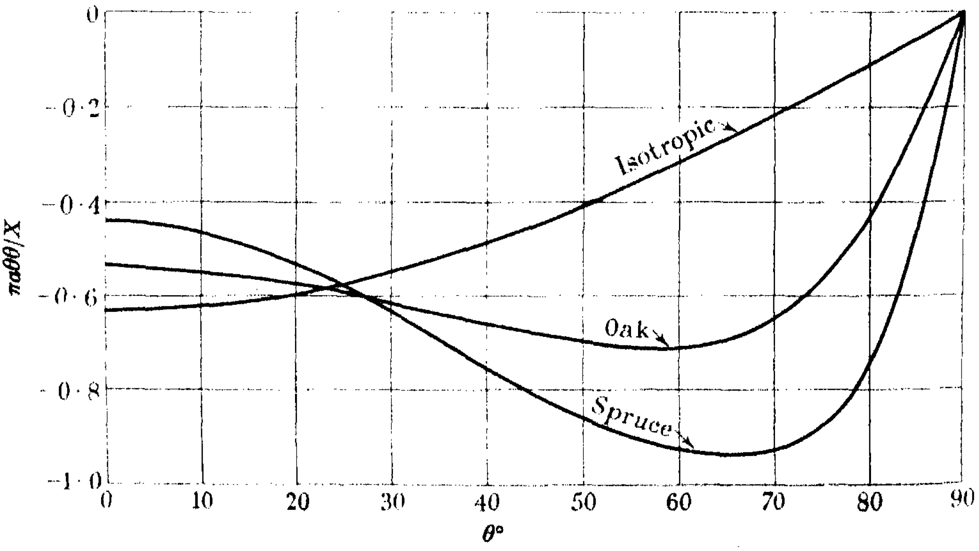


FIGURE 9. Stress distribution round the edge of a hole. Resultant force parallel to the grain and to $\theta = 0$.

force is perpendicular to the grain, the resultant force always being parallel to the line $\theta = 0°$. The distribution of $\widehat{\theta\theta}$ for a hole in an isotropic sheet for which Poisson's ratio is 0·25 (Bickley 1928) is shown in both figures for comparison.

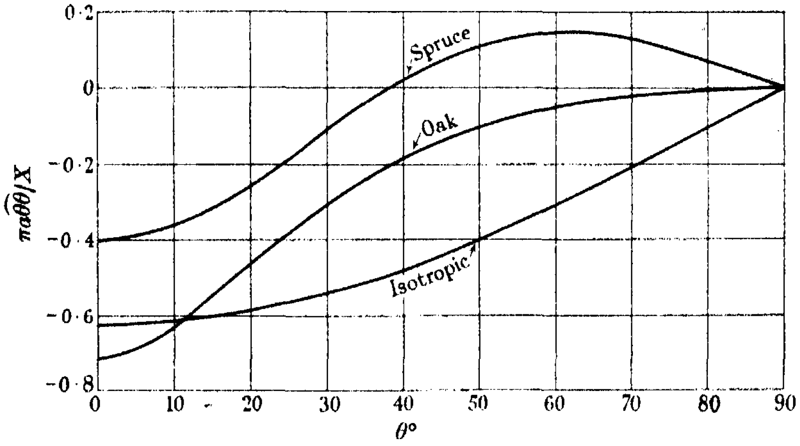


FIGURE 10. Stress distribution round the edge of a hole. Resultant force perpendicular to the grain and parallel to $\theta = 0$.

14. Bickley (1928) attempted to give some idea of the stress distribution round a rivet driven into a hole in an isotropic sheet and pulled sideways, by taking the edge stresses to be given by

$$rr_c = -P(1+\cos\theta), \quad \widehat{r\theta}_c = \tfrac{1}{4}P\sin\theta. \tag{14·1}$$

This system of stresses gives a resultant force $X = \frac{5}{4}\pi a P$ parallel to $\theta = 0$. It is interesting to study the same distribution of stress for aeolotropic plates which may be deduced from §§ 10, 12 and 13. The values of the stresses at the edge of the hole are given in table 8*a* for spruce and table 8*b* for oak, and the distribution $\widehat{\theta\theta}$ round two quadrants of the hole is shown in figure 11 where the resultant force X is along the line $\theta = 0°$. The value of $\widehat{\theta\theta}$ on the edge of a hole in an isotropic material for which Poisson's is 0·25 is shown in figure 11 for comparison. It should be noticed that a correction appears to be necessary to Bickley's results as his values of $\widehat{\theta\theta}$ for the range $\theta = 0° - 90°$ appear to be the ones which should actually be given for the range $\theta = 180° - 90°$, and vice versa. It will be seen that the stress distribution for spruce shows some remarkable differences from that for an isotropic material.

Bickley discussed various other distributions of stress in an isotropic plate containing a circular hole and, if needed, the corresponding results for aeolotropic plates may be deduced from the general formula of § 4.

TABLE 8*a*. VALUES OF STRESSES AT THE EDGE OF THE HOLE: SPRUCE

$\theta°$	$\dfrac{\pi a\widehat{xy}}{X}$	$\dfrac{\pi a\widehat{xx}}{X}$	$\dfrac{\pi a\widehat{yy}}{X}$	$\dfrac{\pi a\widehat{\theta\theta}}{X}$
		Grain parallel to resultant force		
0	0	−1·6	0·924	0·924
10	−0·398	−1·52	0·865	0·929
20	−0·748	−1·30	0·693	0·940
30	−1·01	−0·970	0·423	0·946
40	−1·13	−0·575	0·0841	0·922
50	−1·08	−0·211	−0·281	0·822
60	−0·857	−0·0164	−0·606	0·578
70	−0·522	−0·156	−0·815	0·103
75	−0·365	−0·366	−0·858	−0·216
80	−0·263	−0·567	−0·858	−0·486
85	−0·233	−0·483	−0·832	−0·445
90	−0·200	0·184	−0·8	0·184
95	0·0373	1·12	−0·751	1·10
100	−0·221	1·71	−0·657	1·72
105	−0·460	1·84	−0·522	1·92
110	−0·624	1·71	−0·368	1·86
120	−0·745	1·19	−0·071	1·52
130	−0·685	0·713	0·160	1·16
140	−0·544	0·377	0·308	0·873
150	−0·386	0·173	0·388	0·668
160	−0·240	0·0639	0·422	0·534
170	−0·113	0·0139	0·433	0·459
180	0	0	0·435	0·435
		Grain perpendicular to resultant force		
0	0	−1·6	−0·0254	−0·0254
5	−0·148	−1·59	0·292	0·303
10	−0·400	−1·52	0·878	0·942
15	−0·715	−1·40	1·29	1·47
20	−1·01	−1·21	1·42	1·76
30	−1·41	−0·734	1·13	1·89
40	−1·52	−0·242	0·557	1·73
50	−1·40	0·176	−0·00754	1·48
60	−1·14	0·474	−0·442	1·23
70	−0·814	0·646	−0·709	1·01
80	−0·487	0·707	−0·818	0·827
90	−0·2	0·680	−0·8	0·680
100	−0·0239	0·591	−0·692	0·561
110	−0·174	0·469	−0·531	0·464
120	−0·252	0·337	−0·354	0·383
130	−0·268	0·216	−0·190	0·312
140	−0·239	0·121	−0·0557	0·253
150	−0·187	0·0586	0·0438	0·210
160	−0·134	0·0254	0·132	0·206
165	−0·110	0·0161	0·191	0·235
170	−0·0839	0·00875	0·266	0·287
175	−0·0480	0·00266	0·346	0·352
180	0	0	0·393	0·393

TABLE 8*b*. VALUES OF STRESSES AT THE EDGE OF THE HOLE: OAK

$\theta°$	$\dfrac{\pi a\widehat{xy}}{X}$	$\dfrac{\pi a\widehat{xx}}{X}$	$\dfrac{\pi a\widehat{yy}}{X}$	$\dfrac{\pi a\widehat{\theta\theta}}{X}$
Grain parallel to resultant force				
0	0	−1·6	1·03	1·03
10	−0·412	−1·52	0·948	1·01
20	−0·760	−1·30	0·726	0·978
30	−0·991	−0·978	0·397	0·912
40	−1·07	−0·620	0·0199	0·813
50	−1·01	−0·295	−0·340	0·679
60	−0·832	−0·0583	−0·620	0·522
70	−0·617	0·104	−0·781	0·397
80	−0·420	0·326	−0·830	0·434
90	−0·200	0·744	−0·800	0·744
100	−0·116	1·12	−0·675	1·10
110	−0·418	1·14	−0·443	1·22
120	−0·580	0·905	−0·165	1·14
130	−0·600	0·612	0·0889	0·986
140	−0·523	0·360	0·283	0·830
150	−0·399	0·181	0·411	0·700
160	−0·262	0·0721	0·484	0·605
170	−0·128	0·0166	0·518	0·547
180	0	0	0·529	0·529
Grain perpendicular to resultant force				
0	0	−1·6	0·771	0·771
10	−0·393	−1·53	0·835	0·898
20	−0·807	−1·28	0·853	1·12
30	−1·14	−0·892	0·656	1·26
40	−1·31	−0·426	0·295	1·28
50	−1·28	0·0283	−0·112	1·23
60	−1·10	0·402	−0·466	1·14
70	−0·817	0·655	−0·708	1·02
80	−0·499	0·775	−0·816	0·898
90	−0·2	0·777	−0·8	0·777
100	−0·0415	0·691	−0·688	0·664
110	−0·206	0·555	−0·520	0·562
120	−0·293	0·407	−0·331	0·476
130	−0·317	0·275	−0·148	0·413
140	−0·304	0·175	0·0213	0·384
150	−0·271	0·108	0·191	0·406
160	−0·227	0·0593	0·388	0·496
170	−0·144	0·0193	0·606	0·637
180	0	0	0·718	0·718

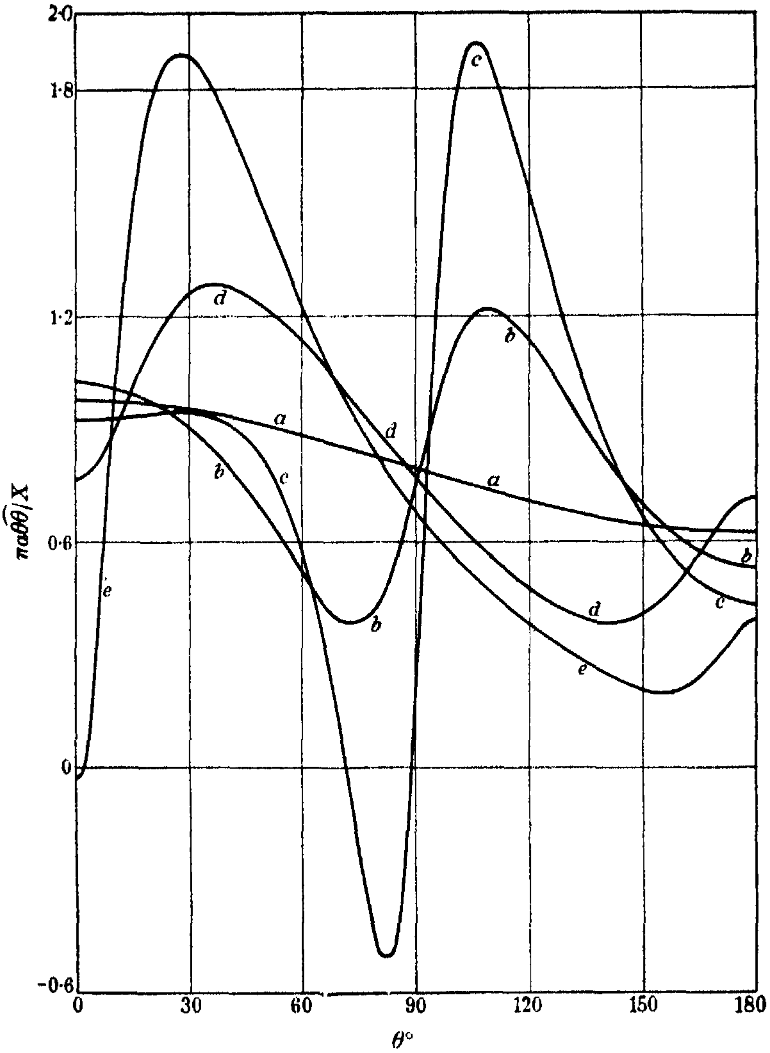


FIGURE 11. Distribution of stress round the edge of a hole. Resultant force parallel to $\theta = 0$. Isotropic material (*a*); grain parallel to resultant force: oak (*b*), spruce (*c*); grain perpendicular to resultant force: oak (*d*), spruce (*e*).

In conclusion the writer wishes to thank the Durham College Research Committee for the use of a calculating machine.

REFERENCES

Bickley, W. G. 1928 *Phil. Trans.* A, **227**, 383.
Green, A. E. 1939 *Proc. Roy. Soc.* A, **173**, 173.
Green, A. E. and Taylor, G. I. 1939 *Proc. Roy. Soc.* A, **173**, 162.
Green, A. E. and Taylor, G. I. 1941 Unpublished.

The effect of a free surface on compressional (shock) waves

By W. G. Bickley, D.Sc.

Imperial College of Science and Technology

(*Communicated by L. Bairstow, F.R.S.—Received* 14 *August* 1941)

The effect of a free surface in releasing the shock pressure is studied, for both shallow and deep water, by the exact mathematical solution of several somewhat idealized problems. Formulae emerge which display discontinuities of the type which could be qualitatively predicted on physical grounds, as representing the effect of successive reflexions at bed and free surface; some graphs showing the quantitative march of events calculated from the formulae are given. The pressure seems never to exceed the shock pressure due to the maximum velocity of impact, but suction occurs in some cases; this suction may have serious effects as regards erosion. The surface elevation shows the characteristic splash at the wall, becoming mathematically infinite there.

Introduction

The present investigation originated as an attempt to provide a theoretical background to certain experiments on wave impact carried out in the hydraulics laboratory of the City and Guilds College. As the experimental results grew in number and were analysed, it became clear that the occlusion of an air bubble was responsible for what were, in the early stages, regarded as anomalous results. Although I am not yet able to include in the mathematical investigation even a highly idealized 'air bubble', such results as have been obtained are of very considerable mathematical and physical interest.

The type of problem considered here is the propagation and release of the impact pressure, and other circumstances of the resulting fluid motion, when moving water having a free surface impinges on a vertical wall. It is actually more convenient to consider a moving wall and (initially) stationary water. Since the speed of compressional waves so greatly exceeds that of surface waves and ripples, gravity and surface tension are negligible; the effects of viscosity are also neglected.

1. Fundamental equations

The first problem may be formulated thus:

Liquid of uniform depth h in a channel of uniform breadth extends to infinity on one side of a vertical cross-channel partition. The partition is constrained suddenly to commence moving with velocity V. It is sought to determine the resulting fluid motion and pressure.

Let x be the distance measured along the channel from the initial location of the partition, and y the distance measured downwards from the equilibrium level of the free surface; let also t denote time measured from the instant at which the partition commences to move. If the viscous effects due to the bed and the walls (in any case small) are neglected, the motion will be independent of the third space co-ordinate. The motion will also be regarded as small, so that squares and higher powers of the velocity components will be neglected.

With these assumptions a velocity potential ϕ' will exist for the motion, in terms of which the velocity components will be given by

$$u_x = -\partial\phi'/\partial x, \quad u_y = -\partial\phi'/\partial y, \tag{1·1}$$

and which satisfies the equation

$$\frac{\partial^2\phi'}{\partial x^2}+\frac{\partial^2\phi'}{\partial y^2}=\frac{1}{c^2}\frac{\partial^2\phi'}{\partial t^2}, \tag{1·2}$$

in which c is the velocity of compressional waves. The excess pressure p is given by

$$p=\rho\frac{\partial\phi'}{\partial t}, \tag{1·3}$$

where ρ is the density of the liquid.

The boundary conditions which must be satisfied are

$$\left.\begin{array}{ll}\text{(i)} & u_x = -\partial\phi'/\partial x = V, \quad x=0,\\ \text{(ii)} & u_y = -\partial\phi'/\partial y = 0, \quad y=h,\\ \text{(iii)} & p = \rho\partial\phi'/\partial t = 0, \quad y=0,\end{array}\right\} \quad (t>0). \tag{1·4}$$

It is convenient to introduce non-dimensional co-ordinates ξ, η, τ, defined by

$$\xi = \pi x/2h, \quad \eta = \pi y/2h, \quad \tau = \pi ct/2h, \tag{1·5}$$

and so chosen that in terms of the new units, the depth of the channel is $\frac{1}{2}\pi$ and the wave velocity unity. The shock pressure 'due to' a velocity V is ρcV; and thus

$$p = \rho cV\varpi, \tag{1·6}$$

where, again, ϖ is non-dimensional. Finally, a non-dimensional velocity potential ϕ is introduced defined by

$$\phi' = (2hV/\pi)\phi. \tag{1·7}$$

In terms of these new variables,

$$u_x = -V\partial\phi/\partial\xi, \quad u_y = -V\partial\phi/\partial\eta, \tag{1·1a}$$

$$\frac{\partial^2\phi}{\partial\xi^2}+\frac{\partial^2\phi}{\partial\eta^2}=\frac{\partial^2\phi}{\partial\tau^2}, \tag{1·2a}$$

$$\varpi = \partial\phi/\partial\tau. \tag{1·3a}$$

2. Solution of the equations

Writing
$$\phi = \Sigma A_m \sin m\eta\, f_m(\xi, \tau), \tag{2·1}$$
where, in the summation (as throughout this paper), m takes all *odd* positive integral values, the boundary condition (ii) is satisfied. Writing γ for $\partial/\partial\tau$, the differential equation for f_m (now a function of ξ but an 'operational form' in τ) is
$$d^2f_m/d\xi^2 = (\gamma^2+m^2)f_m. \tag{2·2}$$

The solution of this corresponding to propagation in the positive ξ direction is
$$f_m = \exp\{-\xi\sqrt{(\gamma^2+m^2)}\}.$$

The 'operational' solution
$$\phi = \Sigma A_m \sin m\eta \exp\{-\xi\sqrt{(\gamma^2+m^2)}\}, \tag{2·3}$$
is thus obtained.

The coefficients A_m must be determined so as to satisfy condition (i); the usual Fourier method yields
$$A_m\sqrt{(\gamma^2+m^2)} = 4/\pi m,$$
so that
$$\phi = \frac{4}{\pi}\Sigma\frac{\sin m\eta}{m}\frac{e^{-\xi\sqrt{(\gamma^2+m^2)}}}{\sqrt{(\gamma^2+m^2)}}. \tag{2·4}$$

To interpret this use is made of the known result
$$\frac{\gamma e^{-\xi\sqrt{(\gamma^2+m^2)}}}{\sqrt{(\gamma^2+m^2)}} = \begin{cases} 0, & 0<\tau<\xi, \\ J_0\{m\sqrt{(\tau^2-\xi^2)}\}, & \tau>\xi. \end{cases}$$

Hence, finally
$$\phi = \frac{4}{\pi}\Sigma\frac{\sin m\eta}{m}\int_\xi^\tau J_0\{m\sqrt{(\alpha^2-\xi^2)}\}\,d\alpha, \tag{2·5}$$

and from this it readily follows that
$$\varpi = \frac{4}{\pi}\Sigma\frac{\sin m\eta}{m}J_0\{m\sqrt{(\tau^2-\xi^2)}\}. \tag{2·6}$$

Formulae for the surface velocity $\dot{s}$ and the surface elevation s respectively are readily derived; they are:
$$\dot{s} = \frac{4}{\pi}V\Sigma\int_\xi^\tau J_0\{m\sqrt{(\alpha^2-\xi^2)}\}\,d\alpha, \tag{2·7}$$
$$s = \frac{4}{\pi}\frac{V}{c}h\Sigma\int_\xi^\tau d\alpha'\int_\xi^{\alpha'} J_0\{m\sqrt{(\alpha^2-\xi^2)}\}\,d\alpha. \tag{2·8}$$

3. SUMMATION OF SERIES AND DISCUSSION OF RESULTS

The above results involve Schlömilch series of Bessel functions, and it is possible to express their sums in finite (albeit discontinuous) form. Watson (*Theory of Bessel functions*, ch. 19) implies that few practical applications of such series have so far been discovered, but it will be found that their mathematical eccentricities are peculiarly appropriate to the present problem.

The series for ϖ will first be summed. Writing, for brevity, $\surd(\tau^2-\xi^2)=\beta$, and using the formula

$$J_0(m\beta)=\frac{2}{\pi}\int_0^{\frac12\pi}\cos(m\beta\sin\theta)\,d\theta,$$

$$\varpi=\frac{8}{\pi^2}\Sigma\frac{\sin m\eta}{m}\int_0^{\frac12\pi}\cos(m\beta\sin\theta)\,d\theta. \tag{3·1}$$

is obtained. Changing the order of summation and integration (easily justified), it is found that

$$\varpi=\frac{4}{\pi^2}\int_0^{\frac12\pi}\Sigma\frac{\sin m(\eta+\beta\sin\theta)+\sin m(\eta-\beta\sin\theta)}{m}\,d\theta. \tag{3·2}$$

Now if

$$\frac{4}{\pi}\Sigma\frac{\sin m\alpha}{m}=f(\alpha),$$

then

$$f(\alpha)=\begin{cases}1 & 2n\pi<\alpha<(2n+1)\pi\\ -1 & (2n+1)\pi<\alpha<(2n+2)\pi,\end{cases}$$

where n is any integer, so that

$$\varpi=\frac{1}{\pi}\int_0^{\frac12\pi}\{f(\eta+\beta\sin\theta)+f(\eta-\beta\sin\theta)\}\,d\theta. \tag{3·3}$$

The discontinuities of $f(\eta+\beta\sin\theta)$ occur where $\sin\theta$ has any of the values

$$(\pi-\eta)/\beta,\quad (2\pi-\eta)/\beta,\quad (3\pi-\eta)/\beta,\quad \ldots,$$

while those of $f(\eta-\beta\sin\theta)$ occur when $\sin\theta$ has any of the values

$$\eta/\beta,\quad (\pi+\eta)/\beta,\quad (2\pi+\eta)/\beta,\quad \ldots.$$

These facts lead to the following formulae:

$$\left.\begin{array}{lll}\varpi=0 & & (\tau<\xi),\\ \tfrac12\pi & & (0<\beta<\eta),\\ \tfrac12\pi-\cos^{-1}(\eta/\beta) & & (\eta<\beta<(\pi-\eta)),\\ \tfrac12\pi-\cos^{-1}(\eta/\beta)-\cos^{-1}(\pi-\eta)/\beta & & ((\pi-\eta)<\beta<(\pi+\eta)),\\ \tfrac12\pi-\cos^{-1}(\eta/\beta)-\cos^{-1}(\pi-\eta)/\beta & & \\ \qquad\qquad+\cos^{-1}(\pi+\eta)/\beta & & ((\pi+\eta)<\beta<(2\pi-\eta)),\end{array}\right\} \tag{3·4}$$

and so on.

Any boundary, $\beta = q\pi \pm \eta$, between regions in which different formulae hold, is the circle

$$\xi^2 + (q\pi \pm \eta)^2 = \tau^2$$

of radius τ and centre $(0, \pm q\pi)$. The diagram (figure 1) exhibits these regions when $\tau = 5\pi/2$, and shows that they correspond to a circular wave originating at the origin of co-ordinates and undergoing successive reflexion at the bed and at the free surface. The sequence of signs $(+ - - + + - - \ldots)$ in (3·4) also corresponds with the facts that on reflexion at the bed, no change of type (from condensation to rarefaction, or vice versa) occurs, whereas on reflexion at the free surface, such a change does occur. Note also that the variables x and t occur in the expression for ϖ only in the combination β; the march of events in thus conveniently exhibited by a graph of ϖ as a function of β for various values of η, such as is given in figure 2. In this diagram the effects of the arrivals of the successive reflected waves are evident.

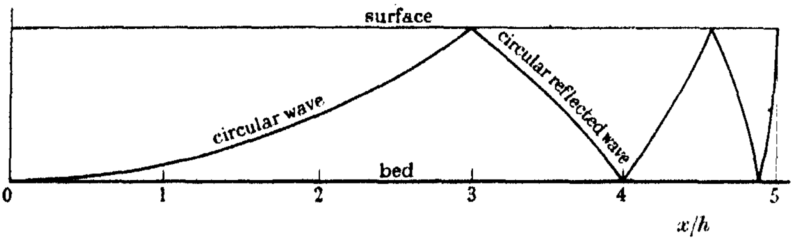


FIGURE 1. Successive wave-fronts.

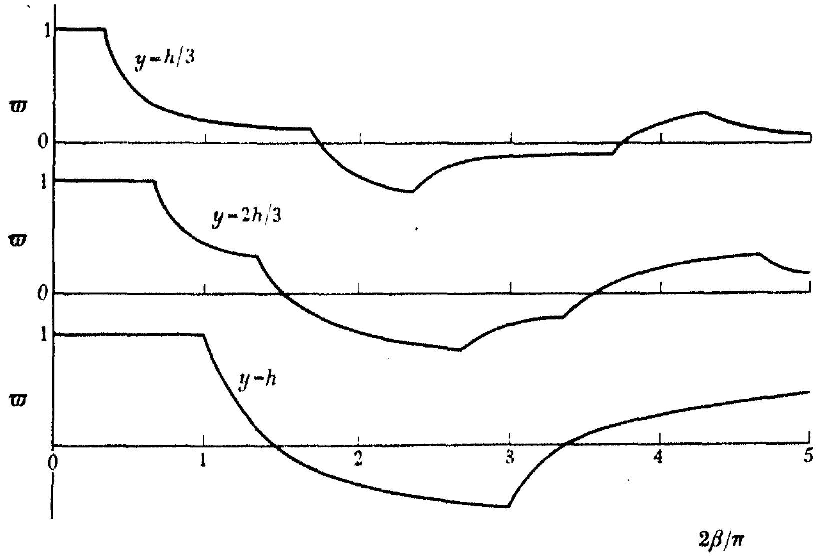


FIGURE 2. Pressure sequence at three depths.

The formulae relating to the motion of the free surface are also of interest. They may be derived from (2·7) and (2·8), or by the aid of (3·4). The former requires first the summation of the series in the integrand of (2·7). Now

$$\begin{aligned}\sum_{m=1}^{2N-1} J_0(m\beta) &= \frac{2}{\pi}\sum_{m=1}^{2N-1}\int_0^{\frac{1}{2}\pi}\cos(m\beta\sin\theta)\,d\theta\\ &= \frac{1}{\pi}\int_0^{\frac{1}{2}\pi}\frac{\sin(2N\beta\sin\theta)}{\sin(\beta\sin\theta)}\,d\theta\\ &= \frac{1}{\pi}\int_0^{\beta}\frac{\sin 2Nu}{\sin u}\,\frac{du}{\sqrt{(\beta^2-u^2)}}.\end{aligned}$$

Since $\sin u = 0$ when $u = 0, \pi, 2\pi, \ldots$, the integral will result mainly from contributions from ranges in the neighbourhood of such values of u. It is known that, if $0 < \alpha < \pi$, and if $f(u)$ is continuous,

$$\lim_{N\to\infty}\int_0^{\alpha}\frac{\sin 2Nu}{\sin u}f(u)\,du = \frac{\pi}{2}f(0),$$

and that, similarly,

$$\lim_{N\to\infty}\int_{q\pi-\alpha}^{q\pi+\alpha}\frac{\sin 2Nu}{\sin u}f(u)\,du = (-)^q\pi f(q\pi).$$

Hence $$\Sigma J_0(m\beta) = \frac{1}{2}\left\{\frac{1}{\beta}-\frac{2}{\sqrt{(\beta^2-\pi^2)}}+\frac{2}{\sqrt{(\beta^2-4\pi^2)}}-\ldots\right\}, \tag{3·5}$$

the series terminating with the term $(-)^r 2/\sqrt{(\beta^2-r^2\pi^2)}$, where r is the integral part of β/π.

This result is confirmed by the use of the relation

$$\frac{\partial^2 s}{\partial\tau^2} = \frac{2}{\pi}\frac{V}{c}h\left(\frac{\partial\varpi}{\partial\eta}\right)_{\eta=0},$$

and (3·4). The results obtained are therefore

$$\dot{s} = \frac{4}{\pi}V\sum_{q=0}^{r}{}'(-)^q\cosh^{-1}\frac{\tau}{\sqrt{(q^2\pi^2+\xi^2)}}, \tag{3·6}$$

$$s = \frac{8}{\pi^2}\frac{V}{c}h\sum_{q=0}^{r}{}'(-)^q\left\{\tau\cosh^{-1}\frac{\tau}{\sqrt{(q^2\pi^2+\xi^2)}}-\sqrt{(\tau^2-q^2\pi^2-\xi^2)}\right\}, \tag{3·7}$$

where the prime indicates that the first term ($q = 0$) in the sum is to be halved. At $\xi = 0$, $\dot{s}$ and s become infinite, so that the assumption of small motion is violated. The graph (figure 3) of s for $\tau = 5\pi/2$ shows, however, that the formula does simulate the characteristic splash (*clapoti*) at $\xi = 0$.

The total force F acting at any instant upon unit width of the partition is given by

$$F = \int_0^h (p)_{x=0}\, dy = \frac{2}{\pi} \rho c V h \int_0^{\frac{1}{2}\pi} \varpi \, d\eta.$$

From (2·6)
$$F = \frac{8\rho c V h}{\pi^2} \Sigma \frac{J_0(m\tau)}{m^2} \tag{3·8}$$

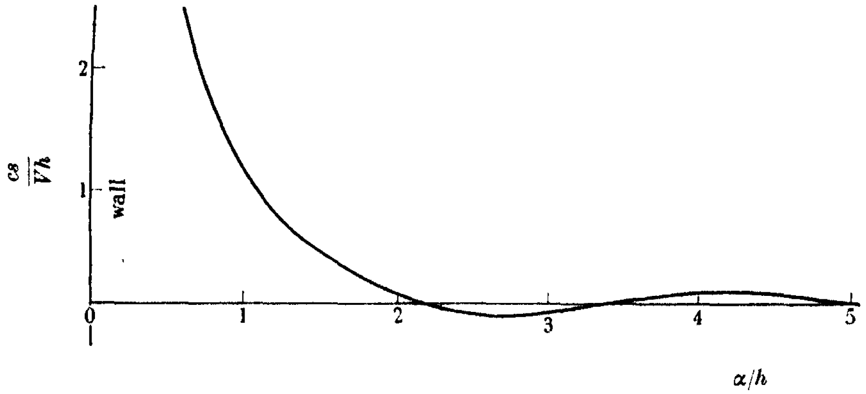


FIGURE 3. Surface elevation when $t = 5h/c$.

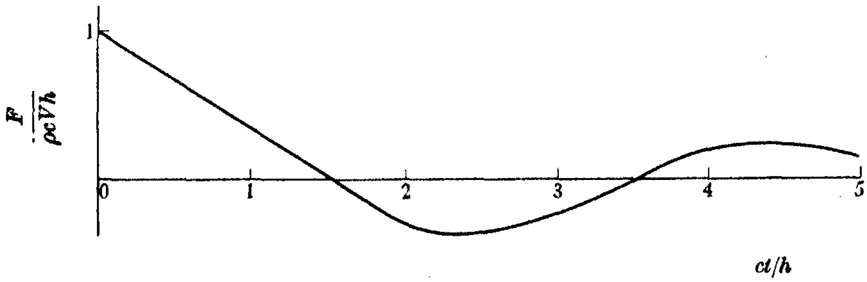


FIGURE 4. Time-variation of total thrust on wall.

is derived. This series can be summed (cf. Watson, *Theory of Bessel functions*, p. 634). Thus, or alternatively after some tedious calculations based on (3·4), it is found that

$$F/\rho c V h = 1 - \frac{4}{\pi^2}[\tau - 2\{\sqrt{(\tau^2 - \pi^2)} - \pi \cos^{-1}(\pi/\tau)\} \\ + 2\{\sqrt{(\tau^2 - 4\pi^2)} - 2\pi \cos^{-1}(2\pi/\tau)\} \ldots], \tag{3·9}$$

the series terminating when succeeding terms would be imaginary. The graph of F as a function of τ (figure 4) again shows, but less markedly, the effects of the sequence of reflexions.

This formula makes F take negative values—a consequence, of course, of the negative pressures shown in figure 2. From the point of view of erosion, the effects of this suction may be more serious than those of the direct shock pressure; such suction will certainly tend to remove particles whose adhesion has been weakened by the impact pressure.

4. GENERALIZATION OF RESULTS

It is easy to modify the above analysis so as to include any law of velocity distribution over the partition $\xi = 0$, i.e. to replace condition (1·4) (i) by

$$u_x = -V\partial\phi/\partial\xi = Vv(\eta), \quad \xi = 0, \tag{4·1}$$

where V is now used as a representative velocity and $v(\eta)$ is a non-dimensional function of η, so far defined only in the range $0 < \eta < \frac{1}{2}\pi$. In this range, $v(\eta)$ can be represented by a series of sines of odd multiples of η. If the definition of $v(\eta)$ is extended so that it represents, for all values of η, the sum of this *series*, then $v(\eta)$ is periodic in η with period 2π and, moreover,

$$\left.\begin{aligned} v(\pi-\eta) &= v(\eta), \\ v(-\eta) &= -v(\eta). \end{aligned}\right\} \tag{4·2}$$

Corresponding to (3·3) it follows that

$$\begin{aligned} \varpi &= \tfrac{1}{2}\int_0^{\frac{1}{2}\pi} \{v(\eta+\beta\sin\theta) + v(\eta-\beta\sin\theta)\}\,d\theta \\ &= \tfrac{1}{2}\int_{-\frac{1}{2}\pi}^{\frac{1}{2}\pi} v(\eta+\beta\sin\theta)\,d\theta. \end{aligned} \tag{4·3}$$

This general solution of the two-dimensional wave equation is closely akin to Whittaker's well-known general solution of Laplace's equation in cylindrical co-ordinates.

The detail of the solution for many reasonable assumed laws of velocity distribution is easily derived from (4·3).

A more important extension is one which will include velocities of the partition variable with *time*. Duhamel's integral applied to the solutions obtained in §§ 2 and 3 gives the relevant formulae. The special case of a constant acceleration α is of some importance, and the formulae may be derived by replacing V by α and integrating from 0 to t. In this way it is found, for instance, that

$$\begin{aligned} F/\rho h^2\alpha &= \frac{ct}{h} - \frac{2}{\pi}\left(\frac{ct}{h}\right)^2 \\ &\quad + \frac{8}{\pi}\Sigma(-)^n n^2\left\{\cosh^{-1}\left(\frac{ct}{2nh}\right) + \frac{ct}{2nh}\sqrt{\left(\frac{c^2t^2}{4n^2h^2}-1\right)} - \frac{ct}{2nh}\cos^{-1}\left(\frac{2nh}{ct}\right)\right\}. \end{aligned} \tag{4·4}$$

5. TRANSITION TO INFINITE DEPTH

This transition is easy, but owing to lack of reflexion from the bed, many features of mathematical and physical interest are lacking in the formulae. The results can, of course, also be obtained independently.

For a constant velocity V of the partition, the solution is contained in the early terms of (3·4), (3·6), (3·7), i.e. (h being now regarded as a representative length)

$$\begin{aligned} \varpi &= \tfrac{1}{2}\pi, && (0<\beta<\eta), \\ &= \tfrac{1}{2}\pi - \cos^{-1}(\eta/\beta), && (\eta<\beta); \qquad (5{\cdot}1) \\ \dot{s} &= 0 && (0<\tau<\xi), \\ &= (2V/\pi)\cosh^{-1}(\tau/\xi) && (\xi<\tau); \qquad (5{\cdot}2) \\ s &= (4Vh/\pi^2 c)\{\tau\cosh^{-1}(\tau/\xi) - \sqrt{(\tau^2-\xi^2)}\}, && (\xi<\tau). \qquad (5{\cdot}3) \end{aligned}$$

With any other law of velocity distribution (where V is now a representative velocity)

$$-(\partial\phi/\partial\xi)_{\xi=0} = v(\eta) \quad (0<\eta<\infty),$$

and if $v(\eta)$ is defined for negative values of η so that

$$v(-\eta) = -v(\eta),$$

the solution is $$\varpi = \tfrac{1}{2}\int_0^{\frac{1}{2}\pi}\{v(\eta+\beta\sin\theta)+v(\eta-\beta\sin\theta)\}\,d\theta. \qquad (5{\cdot}4)$$

The results (5·1), (5·2) and (5·3) readily follow from (5·4).

Practical cases, e.g. water of some depth at a vertical sea wall, would involve a velocity which differed from zero for a small depth (comparable with the height of a wave) only. Details of several such cases have been worked out. In the case where

$$\begin{aligned} u_x &= V, \quad x = 0, \quad (0<y<Y), \\ u_x &= 0, \quad x = 0, \quad (y>Y), \end{aligned}$$

the formulae are not difficult to derive, and the discontinuities in form correspond to circular wave fronts originating at (0, 0) and (0, Y) at $t = 0$. There are four subcases, each with four ranges associated with different formulae. I refrain from writing the sixteen, but for the total force per unit width and for the surface velocity and elevation the number of formulae is reduced to two, as follows:

$$\begin{aligned} F/\rho c V Y &= 1 - \frac{2ct}{\pi Y} \quad (ct<Y) \\ &= \frac{2}{\pi}\left\{\sin^{-1}\frac{Y}{ct} - \frac{ct-\sqrt{(c^2t^2-Y^2)}}{Y}\right\} \quad (ct>Y), \qquad (5{\cdot}5) \end{aligned}$$

$$\begin{aligned} \dot{s} &= \frac{2}{\pi}V\cosh^{-1}\frac{ct}{x} \quad (x<ct<\sqrt{(x^2+Y^2)}), \\ &= \frac{1}{\pi}V\left\{2\cosh^{-1}\frac{ct}{x} - \cosh^{-1}\frac{ct}{\sqrt{(x^2+Y^2)}}\right\} \quad (ct>\sqrt{(x^2+Y^2)}), \qquad (5{\cdot}6) \end{aligned}$$

$$s = \frac{2}{\pi} V \left\{ t \cosh^{-1} \frac{ct}{x} - \sqrt{(t^2 - x^2/c^2)} \right\} \quad (x < ct < \sqrt{(x^2 + Y^2)}),$$

$$= \frac{1}{\pi} V \left\{ 2t \cosh^{-1} \frac{ct}{x} - 2\sqrt{(t^2 - x^2/c^2)} - t \cosh^{-1} \frac{ct}{\sqrt{(x^2 + Y^2)}} + \sqrt{(t^2 - (x^2 + Y^2)/c^2)} \right\}$$

$$(ct > \sqrt{(x^2 + Y^2)}). \quad (5\cdot7)$$

Figure 5 shows how F steadily decreases, while the form of the free surface is shown for four values of t in figure 6.

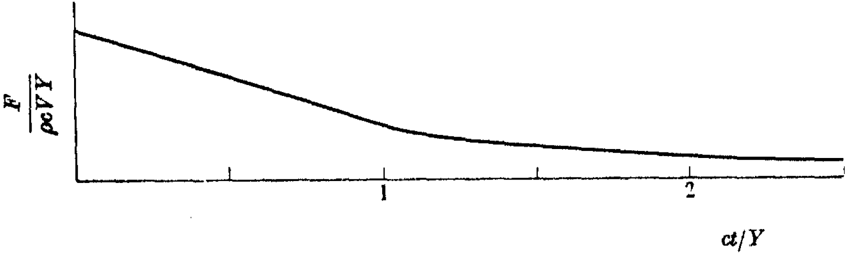


FIGURE 5. Time-variation of total thrust on wall.

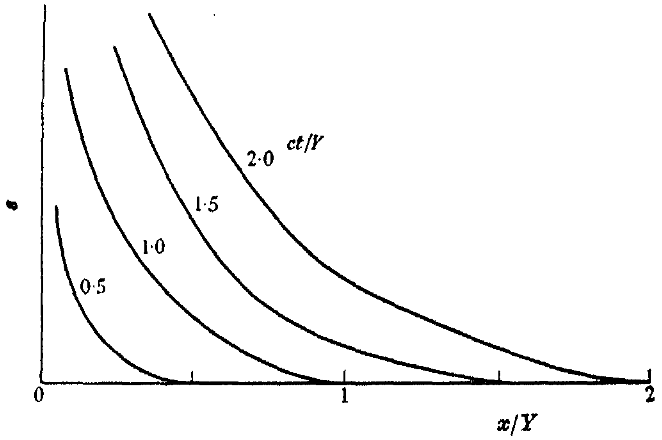


FIGURE 6. A sequence of surface profiles.

[*Note added in proof.* The fact, pointed out by the referee, that the analysis for the case of water of depth h on a rigid bottom is applicable also to a jet of breadth $2h$, is remarkable, since it appears that symmetry alone can mimic the effect of a rigid central partition.]

The 'second maximum' of the shower transition curve of cosmic radiation

By E. P. George, *Birkbeck College, London*,
L. Jánossy and M. McCaig, *University of Manchester*

(*Communicated by P. M. S. Blackett, F.R.S.—Received* 21 *August* 1941)

Various observers have reported, besides the well-established first maximum of the Rossi transition curve, a second maximum for larger thicknesses of absorbers. In the present paper, measurements of the Rossi curve, carried out with various counter systems, are reported. None of these *measurements shows any indication of a second maximum. Using, however*, an arrangement similar to that by Schmeiser & Bothe, a spurious maximum can be obtained under certain circumstances.

Several authors have reported observations indicating that the shower transition curve in lead shows a second flat maximum between 15 and 20 cm. of lead in addition to the well-established first maximum at about 1·5 cm. of lead. More recent extensive observations of other authors did not confirm the previous results. We have collected in table 1 the results put forward by various observers.

The errors given in the last column of table 1 are mostly estimated by us from the published graphs. Owing to the uncertainty introduced by statistical fluctuations, many of the results claiming to show a second maximum would not be incompatible with a curve drawn without a second maximum. On the whole it seems that the transition curve does not show a second maximum, but because of the rather conflicting results quoted in table 1 it appears worth while to report observations on the subject carried out by us. These results are all entirely negative.

1. OBSERVATIONS ON LARGE-ANGLE SHOWERS

(*a*) A variable absorber T was brought close above a triangular threefold coincidence set. Threefold coincidences were observed as a function of the thickness of T. This arrangement is similar to one used by Drigo (1934). The results of the observations are given in table 2.

There is no indication in the figures of table 2 of the occurrence of a second maximum; even allowing for an error due to statistical fluctuations up to three times the standard deviation, no maximum exceeding 7 % of the background is compatible with the observations.

TABLE 1

author	experimental arrangement	position of maximum	magnitude of maximum
Ackemann (1934)	2 counters in a horizontal plane	17 cm. Pb	8 %
Hummel (1934)	,, ,, ,,		20 %*
Drigo (1934)	3 counters in a triangle	17 cm. Pb	(17 ± 5·5) %
Priebsch (1936)	3 counters in a horizontal plane	17 cm. Pb 25 cm. Pb	(6 ± 7) %
Clay, van Gemmert & Wiersma (1936)	3 counters in a triangle	20·5 cm. Pb 30 cm. Fe	(10 ± 7) %
Schwegler (1936)	4 counters in a horizontal plane counting (a) twofold and (b) fourfold coincidences	no maximum	
Woodward (1936)	not stated	18 cm. Pb	9 %*
Morgan & Nielsen (1937)	4 counters in a trapezium	no maximum	
Schmeiser & Bothe (1938)	selecting small angle showers	15 cm. Pb	(25 ± 13) %
Auger, Maze, Ehrenfest & Fréon (1939)	3 counters in a triangle	no maximum	
Nielsen, Morgan & Morgan (1939)	(a) 4 counters in a trapezium	no maximum	
	(b) triangular arrangement	25 cm. Fe	small
Altmann, Walker & Hess (1940)	various arrangements	no maximum	
Swann & Ramsey (1940)	trays of counters controlling electroscopes	30 cm. Sn	small
de Vos (1940)	selecting small angle showers	35 cm. Cu	(10 ± 7) %
Broussard & Graver (1941)†	cloud chamber	17 cm. Pb	(25 ± 15) %

* error not stated. † added to the proof.

TABLE 2

absorber in cm. Pb	0	1·5	5·0	9·0	14·0	16·5	21·5
total count	600	1191	1105	4483	5344	4908	4109
rate per hour	24·4 ± 1·0	33 ± 2·5	45 ± 1·5	39·1 ± 0·6	36·9 ± 0·5	37·2 ± 0·5	36·5 ± 0·6

(b) It was suggested by Drigo (1935) that the showers giving rise to the second maximum were of short range. To investigate this effect a tray of anticoincidence counters was placed below the triangular arrangement. A 1 cm. thick lead absorber was placed between the anticoincidence tray and the coincidence counters (figure 1 a). Thus the anticoincidences observed with this arrangement were mainly due to showers emerging from T and having ranges less than 1 cm. of lead. The observed rates of coincidences and anticoincidences obtained with this arrangement are collected in table 3.

TABLE 3

absorber in cm. Pb:	0	1·5	5	10	17
coincidences per hour	45 ± 3	122 ± 5	60 ± 5	55·5 ± 0·7	52·4 ± 0·7
anticoincidences per hour	22 ± 3	33 ± 3	20 ± 3	17·3 ± 0·5	16·7 ± 0·6

According to table 3 the rates for $T = 17$ cm. of lead are if anything lower than those for $T = 10$ cm. of lead; thus no maximum is indicated.

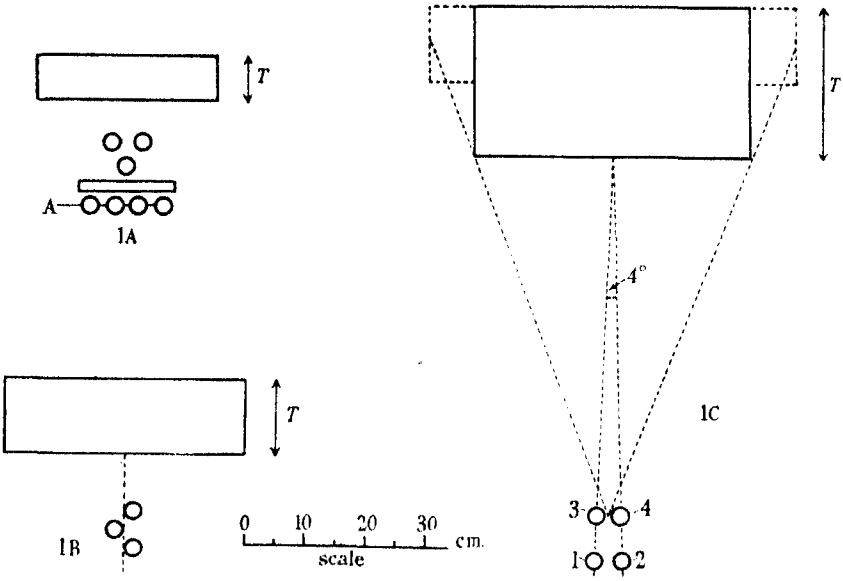


FIGURE 1. Experimental arrangements.

2. OBSERVATIONS ON SMALL-ANGLE SHOWERS

(*a*) According to Schmeiser & Bothe (1938), only the small-angle showers show the second maximum. To investigate small-angle showers we used the arrangement reproduced in figure 1 *b* which appears to be sensitive for small-angle showers. The entirely negative result of the observations obtained with this arrangement are given in table 4.

TABLE 4

absorber in cm. Pb	0	2	5	7	10	12
total count	978	3396	1665	1127	8337	10785
rate per hour	34·5	160·5	69·3	58·6	52·7	50·0
	± 1·1	± 2·8	± 1·7	± 1·7	± 0·8	± 0·5
absorber in cm. Pb	15	17	20	25	30	35
total count	15274	12998	8691	6452	4962	4143
rate per hour	50·1	50·4	49·4	47·6	47·7	46·5
	± 0·4	± 0·4	± 0·5	± 0·6	± 0·7	± 0·7

(*b*) Observations were carried out with the counter arrangement reproduced in figure 1 *c*. This arrangement was very similar to that used by Schmeiser & Bothe. Fourfold coincidences 1–2–3–4 and threefold coincidences 1–2–(3 or 4) were recorded simultaneously. The absorber T was varied in thickness but had a constant cross-section of 46 × 46 cm.^{2}. Neither the threefold rates nor the fourfold rates so observed showed any indication of a second maximum, as can be seen from the data collected in table 5.

TABLE 5

absorber in cm. Pb	0	0·7	1·05	1·4	1·75	5·0
threefolds:						
total count	3820	1326	2200	2791	1235	6288
rate per hour	41·4 ± 0·7	56·0 ± 1·5	56·0 ± 1·2	58·4 ± 1·1	53·6 ± 1·5	43·0 ± 0·5
fourfolds:						
total count	568	309	507	727	309	1407
rate per hour	6·2 ± 0·3	13·0 ± 0·7	12·9 ± 0·6	16·6 ± 0·6	13·4 ± 0·8	8·3 ± 0·2
absorber in cm. Pb	10·0	12·5	15·0	17·5	20·0	25·0
threefolds:						
total count	4759	6002	5486	5698	6812	4554
rate per hour	39·6 ± 0·6	40·6 ± 0·5	39·6 ± 0·5	40·3 ± 0·5	40·3 ± 0·5	40·2 ± 0·6
fourfolds:						
total count	881	1252	1248	1170	1102	734
rate per hour	6·3 ± 0·2	6·4 ± 0·2	6·6 ± 0·2	6·5 ± 0·2	6·5 ± 0·2	6·5 ± 0·2

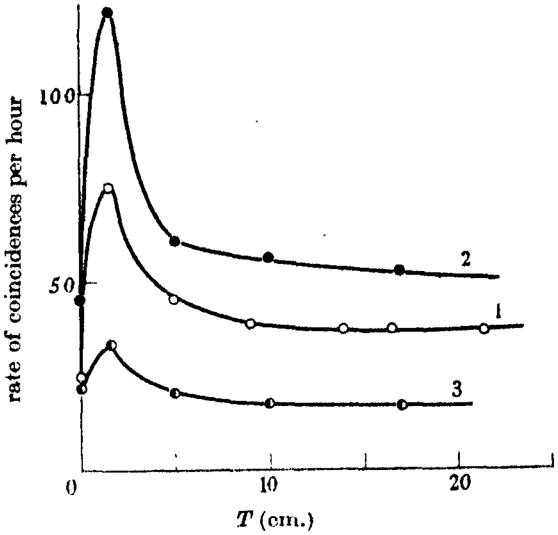


FIGURE 2. Transition effect of large-angle showers. 1, triple coincidences; 2, triple coincidences (absorbers placed below the coincidence counters); 3, anticoincidences.

(c) Another set of observations was carried out with the counters arranged as under (b), but with the absorber modified as follows. The absorber consists of two parts. The lower part is 12·5 cm. thick and 46 × 46 cm.^{2} in cross-section. The upper part varies in thickness but has a cross-section of

61 × 61 cm.². The shape of this absorber is indicated in figure 1c by a broken line. It is rather similar to that used by Schmeiser & Bothe, although the exact areas of their absorbers are not always clearly stated.

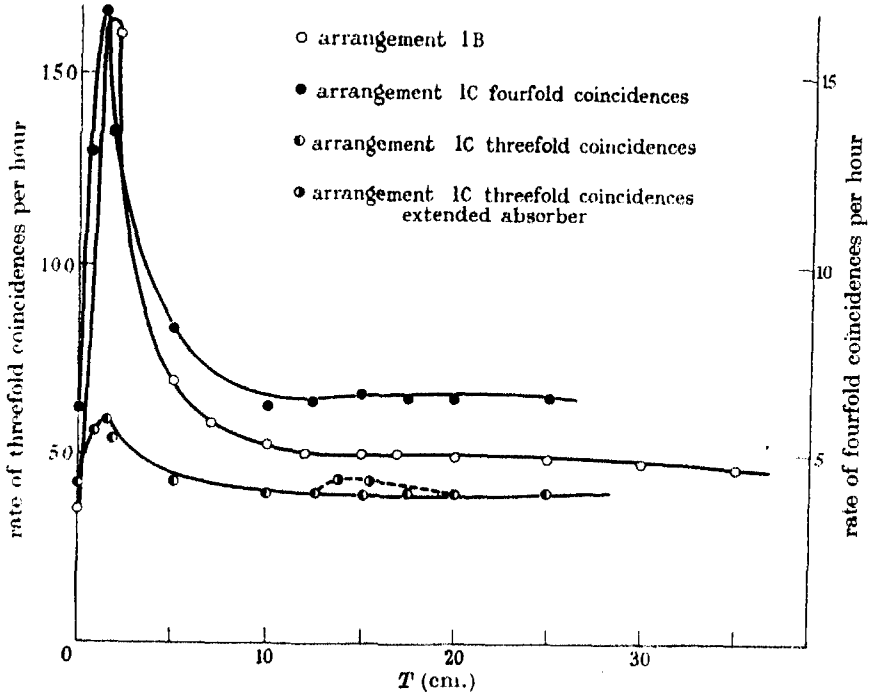


FIGURE 3. The transition effect of narrow-angle showers.

The observations obtained with the above arrangement are given in table 6.

TABLE 6

thickness of absorber in cm. Pb	12·5	12·5 ± 1·4	12·5 ± 2·8	12·5 ± 7·4
threefold coincidences	40·6 ± 0·5	43·5 ± 0·8	42·7 ± 1·0	40·5 ± 0·7
rates per hour				
fourfold coincidences	6·4 ± 0·2	7·1 ± 0·3	6·8 ± 0·4	6·0 ± 0·3

The figures given in table 6 indicate an increase of the shower rate for absorbers thicker than 12·5 cm. Pb. But this increase is clearly spurious. The maximum so obtained is simply the ordinary first maximum produced by the protruding edges of the upper part of the absorber.

This spurious maximum is not observed if the area of the upper part of the absorber is only increased to 54 × 54 cm. instead of 61 × 61 cm. With an area of 54 × 54 cm. the upper part of the absorber subtends approximately the same solid angle at a point in the counter system as is subtended by the lower part of the absorber.

We are therefore led to conclude that the shower transition curve does not show for either small-angle showers or large-angle showers any indication of a second maximum. Moreover, no second maximum exceeding a few percent of the shower intensity would be compatible with our observations.

Some of the observations were carried out in collaboration with Professor B. Rossi and were referred to some time ago (Blackett & Rossi 1939).

REFERENCES

Ackemann 1934 *Naturwissenschaften*, **22**, 169.
Altmann, Walker & Hess 1940 *Phys. Rev.* **58**, 1011.
Auger, Maze, Ehrenfest & Fréon 1939 *J. Phys. Radium*, **10**, 1.
Blackett & Rossi 1939 *Rev. Mod. Phys.* **11**, 281.
Broussard & Graver 1941 *Phys. Rev.* **60**, 413.
Clay, van Gemmert & Wiersma 1936 *Physica, Eindhoven*, **3**, 627.
Drigo 1934 *Ric. Scient.* **5**, 88.
Drigo 1935 *Ric. Scient.* **6**, 529.
Hummel 1934 *Naturwissenschaften*, **22**, 170.
Morgan & Nielsen 1937 *Phys. Rev.* **52**, 564.
Nielsen 1938 *J. Franklin Inst.* **226**, 601.
Nielsen, Morgan & Morgan 1939 *Phys. Rev.* **55**, 995.
Priebsch 1936 *Wien. Ber.* **145**, 101.
Schmeiser & Bothe 1938 *Ann. Phys., Lpz.*, **32**, 161.
Schwegler 1936 *Z. Phys.* **101**, 93.
Swann & Ramsey 1940 *Phys. Rev.* **58**, 477.
de Vos 1940 *Nature, Lond.*, **145**, 387.
Woodward 1936 *Phys. Rev.* **49**, 711.

The distribution of autelectronic emission from single crystal metal points

II. The adsorption, migration and evaporation of thorium, barium and sodium on tungsten and molybdenum

By M. Benjamin, Ph.D. and R. O. Jenkins, Ph.D.*

Communication from the Staff of the Research Laboratories of The General Electric Company Limited, England

(*Communicated by Professor R. H. Fowler, F.R.S.—Received* 30 *July* 1941)

[Plates 2–11]

The adsorption, migration and evaporation of thorium, barium and sodium on tungsten and molybdenum have been studied by means of the autelectronic patterns from single crystal points of these metals. It is found that migration occurs at relatively low temperatures, and that for an optimum thermionic covering there is a unique distribution of the migrating atoms on the metal crystal. Certain crystal zones are avoided by the migrating atoms. The atoms evaporate more readily from certain crystal zones, and the differences in the temperatures for complete migration and evaporation of the atoms lead to an explanation of why migration is not always observed with barium (where the difference between migration and evaporation energy is small) and is always observed with thorium (where the differences are much greater) on multicrystalline surfaces.

In the studies of thorium on tungsten and thoriated tungsten it is shown that diffusion of thorium occurs through the crystal lattice along certain preferred directions and that the final distribution of thorium is identical with that obtained by depositing thorium from an external source.

1. Introduction

In the first part of this paper, a description of the field emission patterns from clean tungsten and molybdenum was given. The work has been extended to a study of the field emission patterns obtained when the clean metal surface is covered with a thin film of electropositive material (thorium, barium, sodium). The surface migration and surface distribution of the contaminating electropositive materials have been investigated. It is found that the deposited electropositive atoms distribute themselves only over certain crystal faces for an optimum covering. The conditions which may inhibit surface migration are also discussed.

---

* I regret to report the death of my collaborator Dr Benjamin in a flying accident on 1 September, 1941.—R.O.J.

2. EXPERIMENTAL PROCEDURE

(*a*) *The tube*

The tube used is similar to that described in the first paper (Benjamin and Jenkins 1940). The source of electropositive material is situated in a small side tube attached to the main bulb, as shown in figure 1. The arrangement is such that the electropositive material deposits itself over one-half of the point surface when evaporated from the source.

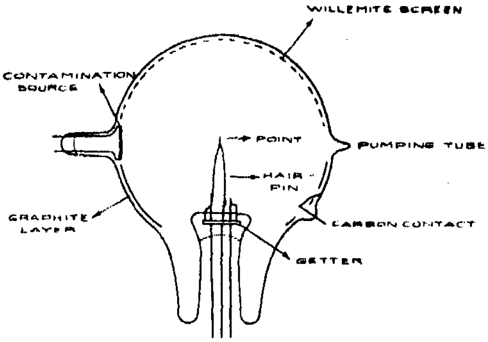


FIGURE 1. Diagram of experimental tube.

(*b*) *The thorium source*

Thorium was obtained in two ways. It was evaporated from a bead of thorium melted on to a tungsten wire in the side tube, or produced by using a point of thoriated tungsten and allowing thorium to diffuse to the surface. The experiment using the thoriated tungsten point we shall refer to as 'thoriated tungsten', and the experiment using thorium deposited from an external source we shall refer to as 'thorium deposited on tungsten'.

(*c*) *The barium source*

The barium source consisted of a strip of tantalum covered with barium beryllate, and is similar to the barium getter source used in each tube. It has already been described in detail (Benjamin and Jenkins 1940).

(*d*) *The sodium source*

The sodium source consisted of a layer of sodium carbonate on a tantalum strip. The carbonate is first decomposed by heating the strip, and on then raising the temperature still further, the sodium oxide is reduced by the tantalum to form the stable tantalum oxide, and sodium which evaporates off the strip.

(*e*) *Operational details*

The tube after assembly was exhausted in the manner previously described, and the source of electropositive material thoroughly outgassed by heating it just below the temperature at which evaporation of the electropositive metal would occur. The tube, after the point and getter heat treatment, was then sealed off, and the point put into the smoothest condition by strongly heating it in the absence of any external electrical field. The point was then lowered to room temperature, and the field emission pattern examined. When it was established that the typical 'smooth' point pattern was being obtained, a thin layer of electropositive metal was deposited over half the point surface by heating the source under controlled conditions. The pattern was then re-examined at room temperature after each heat treatment of the point. All photographs shown were taken at room temperature unless otherwise stated. We would state that many of the higher temperature patterns cannot be photographed when surface mobility is very fast, nor can the appearance be adequately described.

3. THORIUM ON TUNGSTEN

(*a*) *Thoriated tungsten*

The point was prepared from a single crystal Pintsch wire containing 0·8 % thoria.

The point was strongly heated at 2800° K off the pump. At this temperature any thorium produced inside the surface diffuses to and evaporates from the surface, so that on reducing the point temperature to room temperature, the field emission pattern was typical of clean tungsten. The pattern is shown in figure 2*a*. It will be noted that symmetry is about the (100) direction for this wire, instead of the usual (110) direction for all our pure tungsten wires, indicating that in the Pintsch wire the (100) plane lies normal to the wire axis. The (110) and (111) directions can just be discerned. The point was then heated at 2600° K for a few minutes. At this temperature thorium is produced by reduction. The temperature was then reduced to 2000° K to allow thorium to diffuse to the surface. The pattern was viewed at room temperature at frequent intervals. After 3 min. at 2000° K thorium atoms can be seen appearing on the surface, and the applied field necessary to produce a pattern at room temperature is now lower than for the clean point. Figure 2*b* shows the pattern obtained, and it will be noted that the thorium appears to be diffusing to the surface most strongly along the (311) direction, about the (100) centre and the (111) direction. On further heating, the thorium moves off the (100) centre and seems to concentrate over the regions about the (100) (the (311) directions) and over the (111) regions (figure 2*c*). With further heating the thorium spreads over the surface, and distributes itself in a most remarkable manner over the surface, as shown in figure 2*d*. After 30 min. the pattern is that shown in figure 2*e*. The point temperature was then lowered to 1800° K, the usual operating temperature, and after 2 hr. the appearance was as in figure 2*f*. After 19 hr. the appearance was unchanged. Note the distinctive way in which the thorium has spread itself over the surface to form what must be the optimum surface covering. No thorium is present on the (100) centre, but it appears as a series of rings about regions corresponding to the eight (421) planes round the (100) region. The (110) and (211) centres are quite dark, the former being outlined by a ring of thorium. The (111) centre also appears dark, although thorium appears about the (111) region. The pattern is slightly temperature dependent, and in figure 2*g* is shown the appearance of the point in the state of figure 2*f*, but with the point at 1300° K. The bright spot in the (100) centre is the hot tungsten point. The (100), (211), (110) and (111) centres appear less dark than in figure 2*f*, and there is a general haziness which may be due to rapid movement of thorium across the surface, but the regions devoid of thorium are well seen as dark bands. On strongly heating the point at 2400° K, the thorium evaporates quickly from the surface, and some redistribution takes place. The appearance of the surface after 30 sec. at 2400° K is similar to that shown in figure 2*b*, and may be in part due to more thorium diffusing from the interior after the surface thorium has evaporated. The pattern (figure 2*f*) can be obtained again by reactivation at 1900° K.

(*b*) *Thorium deposited on tungsten*

Figure 3 (*a*–*h*) shows the patterns obtained from a pure tungsten point which has had thorium deposited over half its surface from an external thorium source, and has then been progressively heated. Figure 3*a* shows the clean tungsten surface, and figure 3*b* shows the deposit of thorium over half the point surface. On heating the point gently, migration is clearly seen taking place at 870° K, and figure 3*c* shows the pattern obtained after 1 min. at 870° K. The thorium atoms have spread across the surface, avoiding the (110) region, and they have also bunched together to form crystallites where the deposit was thickest. Migration proceeds slowly at 870° K, and is complete within 10 min. Figure 3*d* shows the completely covered surface. The thorium has distributed itself over the surface, avoiding the (110) region and has moved off the (100) and (211) centres. The (111) regions are the most intense emitting zones. Further heating results in redistribution of thorium, and presumably evaporation also occurs until the optimum covering is reached. Figure 3*e* shows the appearance, after 5 min. at 1270° K. Thorium has spread over the (100) zones, and is concentrated at the (211) centres. No thorium has yet appeared close to the (110) direction. Figure 3*f* shows the appearance after a further 5 min. at 1370° K. There is now a marked resemblance to figure 2*f*. The thorium has arranged itself in eight groups about the (100) centre, and there is a ring of thorium around the (110). The (110) and (211) regions are joined by dark bands and the (111) regions are bright, with a faint dark centre.

At this stage the applied field is the lowest recorded, indicating that an optimum value of the surface distribution has been reached. Further heating between 1370 and 1800° K results in very little change in the surface appearance. At 2100° K evaporation is clearly taking place, and the remaining thorium is evenly distributed about the (100) and (111) regions (figure 3*g*). Note here that the (111) region is still brighter than the (100). The surface covering is very sparse and individual atoms or groups can just be resolved. At 2300° K evaporation is almost complete, and the surface is clean except for a few thorium atoms scattered about the (110) zone. The pattern shows now the (100) zone as the most intense emitting, and is very similar to that from completely clean tungsten (figure 3*h*). The rising value of the applied field necessary is due to the rising surface work function as the thorium evaporates.

4. THORIUM ON MOLYBDENUM

An experiment similar to that described in § 3 (*b*) was carried out using a molybdenum point. The autelectronic patterns obtained are shown in figure 4. Figure 4*a* shows the clean molybdenum surface, figure 4*b* the initial deposit of thorium, and figure 4*c* a heavier deposit; figure 4*d* shows migration occurring at 870° K after 2 min. running. Figure 4*e* shows the appearance after 5 min. at 870° K and 5 min. at 1000° K, when the thorium has spread right across the surface, being most intense over the (111). Above 1100° K the surface becomes unstable and gives the impression that over most of the metal surface the thorium is either alloying with the molybdenum or diffusing into the molybdenum interior. Figure 4*f* shows the type of pattern obtained after heating for 5 min. at 1300° K. The thorium has concentrated itself about the (100) and (111) regions, and its resemblance to figure 2*f* is very marked. After a further 5 min. at 1400° K the pattern obtained was that shown in figure 4*g*, when all that can be clearly seen are the eight bright groups about the (100) centre, corresponding as before to the (421) directions, and a brightness on the edge of the (111) zone. The rest of the pattern is dark, probably because of inward diffusion into the molybdenum. The resemblance to figures 2*f* and 3*e* is most clearly seen. Further heating at 1700° K results in some evaporation or inward diffusion, while the grouping of thorium about the (100) in the (421) direction, still persists. Above this temperature complete evaporation or inward diffusion is quite rapid.

5. BARIUM ON TUNGSTEN

The migration of barium on a tungsten point has been studied by Muller (1938), in which he showed that the barium migrated rapidly across the point surface at a temperature of 700° K. We wish to describe two experiments with barium on tungsten, one in which migration was observed and one in which no migration occurred.

(a) *Migration of barium (thin deposit)*

The experiment was similar to that for thorium on tungsten. Figure 5 shows the patterns obtained at various stages of heating of the point after barium deposition. Figure 5*a* is the typical clean tungsten pattern; figure 5*b* shows a thin deposit of barium over half the point. On raising the point to 600° K marked oscillatory motion, but no rapid flow of barium, is observed. Flow becomes evident at 700° K, and at 770° K migration is well advanced in 30 sec. Figure 5*c* shows the spread of barium, and in figure 5*d* the surface is completely covered after 1 min. at 850° K. Note here that the barium in moving across the surface avoids the (100) zone and the centre of the (110) zone. Note also the ring configuration in the (100) region as for thorium. Figure 5*d* represents the nearest approach to an optimum covering. On raising the temperature to 900° K the barium concentrates itself about the (211) region (figure 5*e*). Further raising of the temperature results in evaporation of barium. Figure 5*f* shows the surface covered very sparsely with barium atoms after heating at 1000° K for 5 min. Here only the regions about the (111) centre and between the (110) and (211) centres are covered, showing that evaporation has taken place from the (211) centres. At 1200° K very little barium is left on the surface, and the groups of barium atoms can be clearly seen grouped over the (111) zone and on the regions joining the (100), (110) and (211) zones (figure 5*g*). If the pattern is viewed at the high temperature, the barium atoms can be seen streaming across the (210) regions rather like falling raindrops, but at room temperature the pattern always 'freezes' without any sign of barium in the (211), (110) or (100) regions. At 1400° K the surface is practically clean (figure 5*h*).

(b) *Migration of barium (thick deposit)*

If the experiment is repeated, using a thick initial deposit of barium, substantially the effects described in (*a*) are observed, but there is a period during which the barium, where it is deposited thickest, forms itself into groups of crystallites. Figure 6 (*a*–*d*) shows the effect. Figure 6*a* shows a heavy deposit of barium, the surface of the point nearest the left of the photograph being so thickly covered that its thermionic activity is very much less than that of the central zone of the surface which has a covering approaching a monatomic layer. On heating the point at 700° K barium spreads over the uncovered part of the surface, rather like a thick film of liquid, and on the thickly covered part of the surface crystallites are formed, which project above the surface and appear as very bright spots, because of the very high local fields (figure 6*b*). At 850° K (figure 6*c*) the surface is completely covered in less than 1 min., and crystallites are now present over the part of the surface most remote from the source of barium. At 900° K (figure 6*d*) these crystallites have disappeared, and the surface is covered with barium in a manner similar to that of figure 5*d*.

(*c*) *The non-migration of barium*

Using the same tube as for the experiments in (*a*) and (*b*), a case of non-migration was observed when the barium source was practically exhausted, and prolonged heating of the source at a higher temperature than usual was necessary in order to obtain a barium deposit over the half surface of the point. Figure 7*a* shows the rather sparse deposit of barium over the tungsten surface. On heating the point no movement of barium was discernible until the point temperature was at 1000° K, when coalescence and movement backwards of the barium edge was observed (figure 7*b*). At 1200° K the break-up of the barium film into bright patches was still further apparent (figure 7*c*), and at 1700° K the appearance was that of figure 7*d*. The surface had to be heated to 1900° K to remove completely the last traces of barium, and then the tungsten surface required a higher applied field in order to obtain the original current from it. To recover the tungsten emission the point needed flashing up to a very high temperature. We infer from these results that the intense heating of the barium source must have resulted in the liberation of an electronegative gas (probably oxygen) with the barium, and the presence of this very small trace of gas was sufficient to inhibit the surface migration of barium. We shall return to this matter in the discussion of the results.

6. Barium on molybdenum

(*a*) *Migration over the smooth surface*

The molybdenum point was brought into its smoothest state by flashing at a high temperature (figure 8*a*), and then a thick layer of barium deposited on its surface, so that the more thickly covered part appeared relatively dark compared with the thinner deposit on the centre zone of the surface (figure 8*b*). Further deposition of barium results in a distinct monatomic edge (figure 8*c*). The edge clearly outlines the (110) dark centre, which appears to be a piece of evidence that for this particular molybdenum surface the (110) centre must be slightly below the main surface. On the other hand, the fact that at this part of the surface the covering is only monatomic may mean that barium atoms present in the surface do not affect any directions of low autelectronic emission, which are still determined by the crystal structure of the underlying metal. These remarks have a bearing on our explanations for the appearance of the patterns from a clean metal surface (Part I, § 10). On heating the point at 600° K migration just commences, and crystallite formation occurs after 1 min. (figure 8*d*). On raising the temperature slowly, migration proceeds and the crystallites tend to break and disappear at 800° K when migration is complete, and the surface shows a concentration of barium around the (100) centres (figure 8*e*). At 850° K the surface appears uniformly covered and all crystallites have disappeared (figure 8*f*). At 900° K (figure 8*g*), evaporation has started and barium leaves the (110) region first. At 950° K barium is still over the (100), but has left the (111) centres, and is evident around the (211) centres (figure 8*h*). At 1200° K the barium has left the (100) and is apparent as small groups of atoms about the (211) centres, and is particularly evident at points corresponding approximately to the (433) directions (figure 8*i*). At 1300° K the surface has changed again, more evaporation has taken place from the (100) zone and around the (211) centres (figure 8*j*), and the surface is about clean. At 1400° K the surface returns to clean molybdenum.

(*b*) *Barium migration on built-up surface*

As we showed in the first paper, the surface of molybdenum can be modified by heating it at a high temperature with field applied. Barium migrating across this surface should retain the outline pattern due to the geometrical shape of the built-up point. In figure 9 we show the deposition and migration across a distorted molybdenum surface, all treatment being made with field applied. Figure 9*a* shows the clean distorted surface, and figure 9*b* the initial deposit which is quite thick. On heating, migration occurs rapidly at 700° K (figure 9*c*) and is complete after several minutes (figure 9*d*). At 850° K evaporation from the (110) and (111) centre which has previously not appeared, is apparent (figure 9*e*) and the (433) directions can be seen clear of barium. At 1300° K the surface is sparsely but fairly uniformly covered over all ridges in the surface (figure 9*f*). Throughout, the 'built-up' surface features have been retained, supporting our explanation that changes of surface geometry are the causes of the pattern changes brought about by heating the clean surface with applied field.

7. SODIUM ON TUNGSTEN AND MOLYBDENUM

Sodium when deposited on a tungsten or molybdenum smooth surface behaves exactly the same on each metal. If the layer is fairly thin, no migration is observed at room temperature but it is apparent at 500° K.

If the layer is thick complete migration will take place at room temperature. To illustrate the effect we show two sets of photographs, one set shows an approximately monatomic sodium film on smooth molybdenum, and the other set a thicker sodium film on built-up molybdenum. Figure 10*a* shows the smooth molybdenum surface. Figure 10*b* shows the initial surface covering of sodium. Figure 10*c* shows some migration of sodium at 300° K. At 500° K migration is proceeding (figure 10*d*), and is almost complete after 3 min. (figure 10*e*). Note in figures 10*d* and *e* that no sodium appears on the (110) zone, but covers the (211) zones. At 600° K evaporation is marked (figure 10*f*), and at 700° K only the (211) centres show much sign of sodium (figure 10*g*). At 850° K (figure 10*h*) the sodium is present only about the (211) centres, and at 950° K the surface is clean. Figure 11*a* shows the built-up molybdenum surface, and figure 11*b* the thick deposit of sodium characterized by the bright monatomic edge. Figure 11*c* is the appearance after 20 sec. at room temperature. A considerable degree of movement has taken place. Figure 11*d* shows the appearance after a further 1 min. at room temperature, migration is almost complete; figure 11*e* shows the surface after heating for 2 min. at 500° K, the surface is completely covered, giving the outline built-up pattern. Note that throughout the series the (433) directions about the bright (111) centres, which appear as dark centres in the clean built-up pattern, remain uncovered, whereas in the clean smooth pattern these directions never appear.

8. DISCUSSION

(*a*) *Thorium on tungsten and molybdenum*

The evidence obtained shows that thorium is first mobile over a single crystal tungsten or molybdenum surface at a temperature of about 870° K. Evaporation is first apparent at about 2100° K. Becker and Brittain (1933) showed that thorium migrated over a multicrystalline tungsten surface at about 1500° K, a result confirmed by Benjamin and Jenkins (1940), who pointed out the rapid migration over certain unidentified crystals compared with the slow migration over others. The evidence shows too that for an optimum covering of thorium there is a definite distribution of thorium atoms which results in a surface about one-half covered with thorium. No thorium ever appears on a part of the surface corresponding to a (110) direction.

The final thorium distribution is unchanged whether the thorium is originally obtained by diffusion from within the crystal or by external deposition on the surface. Evaporation appears to take place uniformly from the covered parts of the surface.

(*b*) *Barium on tungsten and molybdenum*

It has been shown that, in agreement with Muller, barium migrates at low temperatures over tungsten and molybdenum. Migration is first evident at about 770° K. The barium distribution is somewhat similar to the thorium, although the photographs obtained do not show the same fine detail as do the thorium photographs. On evaporation, the barium leaves the (110) and (100) regions first, and adheres strongly to the regions around the (211). Evaporation is first evident at about 900° K.

Becker (1929) first showed that barium migrated over a multicrystalline tungsten wire at a temperature of about 900° K. Benjamin and Jenkins (1938) failed to obtain any clear-cut evidence of barium migration, but Becker and Moore (1940), using the same tube as Benjamin and Jenkins, again showed that barium migrated at 970–1070° K, and showed that evaporation was occurring from some crystals while migration was proceeding over others. Becker has suggested that non-migration may be due to the presence of small amounts of gas evolved with the barium, and we, as a result of the experiment on non-migration (§ 5), agree with this view. It is clear, however, that since a wide difference in migration temperatures exists for the single crystal and multicrystalline surfaces, and since migration on a multicrystalline wire takes place at temperatures where barium is appreciably volatile from certain faces on the single crystal surface, cases may arise where the crystal size and orientation of crystals in a polycrystalline wire may give rise to evaporation before migration can occur. This condition is not likely to occur in the case of thorium, since there exists a wide difference in the observed migration temperature and the evaporation temperature, whether on single or polycrystalline surfaces.

We have obtained some further evidence on the migration of barium by repeating our experiments on migration in a tube similar to the tube used by us (Benjamin and Jenkins 1938) and by Becker. Both single crystal and polycrystalline tungsten wires were used, and the tube was made of the hard glass used in the point experiments, so that thorough outgassing at high temperature was possible. In figure 12*a* is shown the initial deposit of barium on the single crystal wire. Migration was very marked at 850° K, and was complete in 10 min. at 850° K, as shown in figures 12*b* and *c*. It is of interest to note the dark bands or flutings corresponding, we believe, to those planes which appear as dark centres in the field emission patterns and which have a much lower thermionic emission as well as lower field emission. In figure 13*a* is shown the initial deposit on the polycrystalline wire. On heating, no movement was observed until 1000° K when migration was noticed, accompanied by some evaporation (figure 13*b*). Migration was complete in 2 hr. at 1100° K (figure 13*c*) over many crystals. The zigzag markings observed are the draw-marks on the wire surface. The filament was then cleaned and run for several hours at 2800° K in the hope of growing larger crystals. The migration experiment was then repeated. Figure 14*a* shows the initial deposit and movement was first clearly visible at 900° K over certain crystals; figure 14*b* shows the movement after 1½ hr. On raising the temperature to 950° K, movement proceeded rapidly, and in ¾ hr. migration was complete over certain crystals (figure 14*c*). Further heating at 1000° K gave no further migration but some evaporation (figure 14*d*). Thus it would appear that in the case of barium the margin between evaporation and migration energies is small, so that on a polycrystalline wire barium may evaporate before migrating, depending on the types of crystal boundaries the migrating atom may have to cross and the type of crystal face on which migration must occur.

(*c*) *Sodium on tungsten and molybdenum*

Sodium migrates readily over both tungsten and molybdenum. Migration is first observed at room temperature, and in the case of thick films of sodium, migration is complete in a very short time at room temperature. Sodium, like thorium and barium, has a preferential distribution over the surface.

The authors wish to express their thanks to Professor R. H. Fowler for the many helpful discussions.

REFERENCES

Becker, J. A. 1929 *J. Amer. Electrochem. Soc.* **45**, 1531.
Becker, J. A. and Brittain, W. H. 1933 *Phys. Rev.* **53**.
Becker, J. A. and Moore, G. E. 1940 *Phil. Mag.* **29**, 129.
Benjamin, M. and Jenkins, R. O. 1938 *Phil. Mag.* **26**, 1049.
Benjamin, M. and Jenkins, R. O. 1940 *Proc. Roy. Soc.* A, **176**, 262.
Müller, E. W. 1938 *Zeit. für Phys.* **108**, 9-10, 668.

The photosensitivity of visual purple solutions and the scotopic sensitivity of the eye in the ultra-violet

By C. F. Goodeve, F.R.S., R. J. Lythgoe and E. E. Schneider

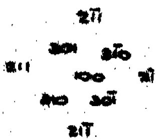
(*Received* 18 *July* 1941)

Previous measurements of the spectral variation of the photosensitivity of visual purple solutions have been extended into the far ultra-violet. The products have been found to be the same as in the visible.

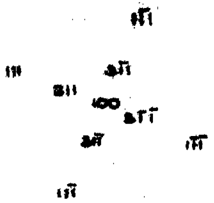
Dark adaptation curves, produced by various observers, including one with an aphakic eye, were of the same type at 365 as at 546 mμ. The limiting scotopic sensitivity (the reciprocal of the potential retinal illumination in quanta/sec./sq. mm. which is just sufficient to excite vision) of normal eyes at 365 mμ is more than ten thousand times smaller than that to be expected from the photosensitivity at this wave-length, a fact which is attributed to the presence of light absorbing substances in front of the retina. The scotopic sensitivity of the aphakic eye was, however, as expected, indicating the absence of absorbing substances. Indeed the aphakic eye was practically as sensitive at 365 as it was at 546 mμ.

The limit of vision in the ultra-violet was accurately determined and, for normal eyes, found to correspond to the threshold of absorption of the lens, 309 mμ. The limit for the aphakic eye (no lens) corresponded to the threshold of absorption of the cornea, 298 mμ.

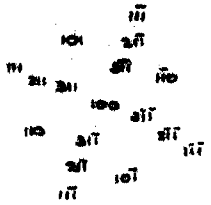
[*This paper has been printed in full in Proceedings B, volume* 130, *pages* 380–395]



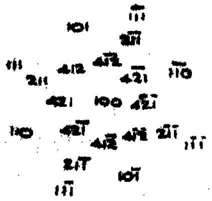
a) HEATED AT 2800 °K
3400 V.
CLEAN TUNGSTEN.



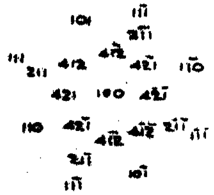
b) HEATED AT 2600 °K, THEN AT 2000 °K 4700 VOLTS
THORIUM DIFFUSING
TO SURFACE.



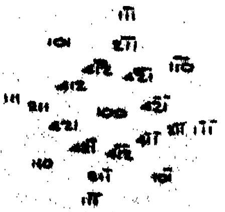
c) FURTHER HEATING AT 2000 °K 4500 VOLTS



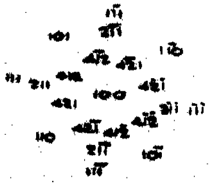
d) AFTER 13 MINUTES AT 2000 °K 4300 VOLTS.



e) AFTER 30 MINUTES AT 2000 °K 4300 VOLTS



f) AFTER 2 HOURS AT 1800 °K APPEARANCE SAME AFTER 10 HOURS. 4000V SURFACE COVERED WITH OPTIMUM AMOUNT OF THORIUM.



g) APPEARANCE OF f ON RAISING POINT TO 1300 °K SHOWS SLIGHT TEMPERATURE DEPENDENCE OF PATTERN.

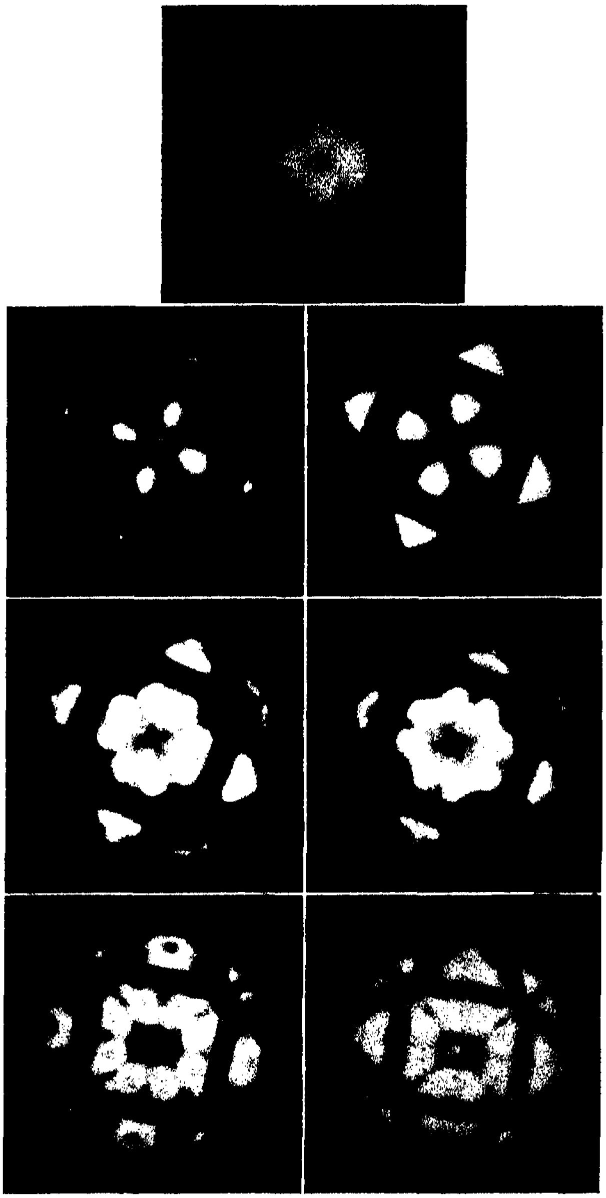


Figure 2. Thoriated tungsten

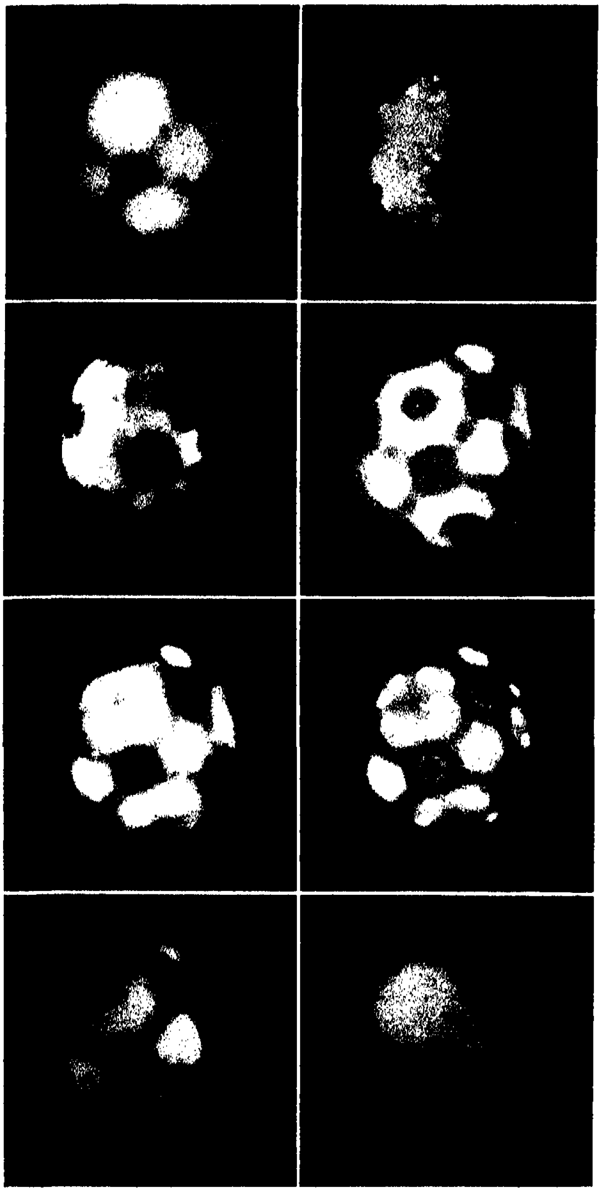
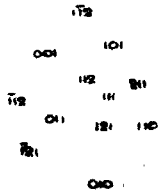
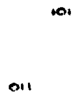


FIGURE 3. Thorium on tungsten.

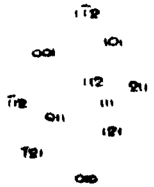


a) CLEAN TUNGSTEN
7300 V.

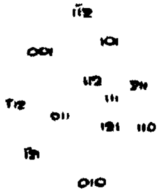
b) THORIUM DEPOSITED
OVER HALF THE POINT
SURFACE 5000 V



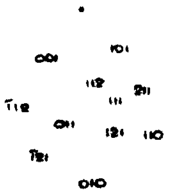
c) POINT HEATED AT 870°K
MIGRATION OCCURS
5000 V.



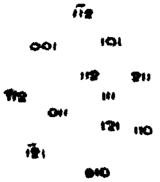
d) AFTER 10 MINUTES AT
870°K MIGRATION
COMPLETE 5000 V



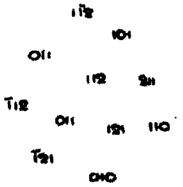
e) HEATED AT 1270°K FOR
5 MINUTES. REDISTRIBUTION
OCCURS 4800V



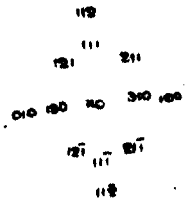
f) 5 MINUTES AT 1370°K
APPEARANCE SIMILAR
TO 2 4450 V



5) AFTER PROGRESSIVELY
RAISING TEMPERATURE
TO 2100 °K WHEN
EVAPORATION OCCURS
5400 V



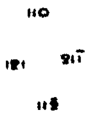
h) HEATED AT 2300 °K
EVAPORATION ALMOST
COMPLETE 7300 V.



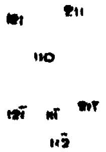
a) CLEAN MOLYBDENUM
18,000 V.

b) INITIAL DEPOSIT OF
THORIUM 15,000 VOLTS.

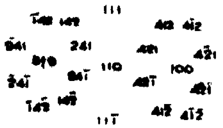
c) HEAVIER THORIUM
DEPOSIT 12,000 VOLTS.



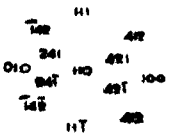
d) MIGRATION OCCURING
AT 870 °K 12,000 VOLTS.



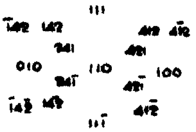
e) MIGRATION AT 870 °K
AND 1000 °K. 11,000 VOLTS.



f) AT 1100 °K — 1300 °K
ALLOYING OR INWARD
DIFFUSION OCCURS
10,000 VOLTS.



g) HEATED AT 1400 °K -
5 MINUTES DISTRIBUTION
SIMILAR TO 2 f
10,000 VOLTS.



h) HEATED AT 1700 °K
EVAPORATION OR INWARD
DIFFUSION OCCURS
13,000 VOLTS.

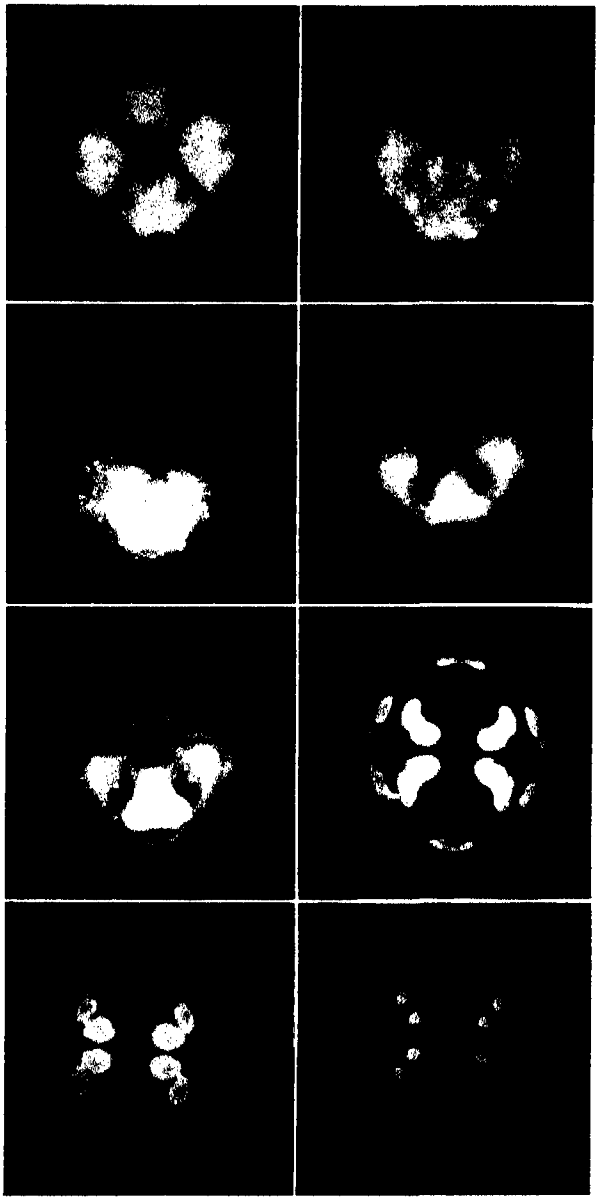


FIGURE 4. Thorium on molybdenum.

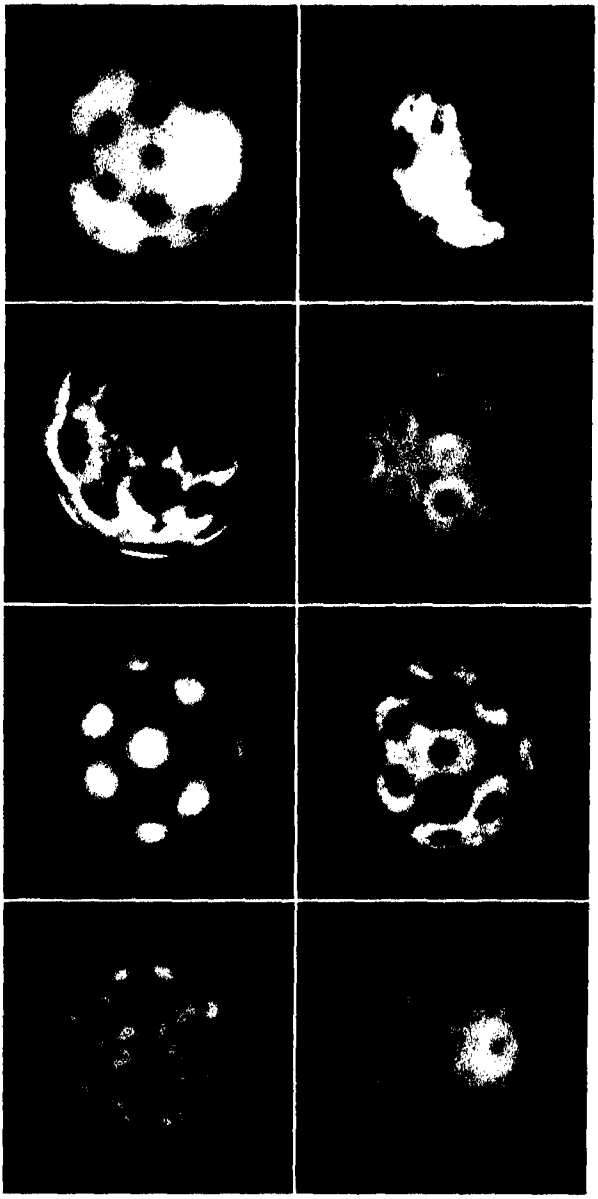
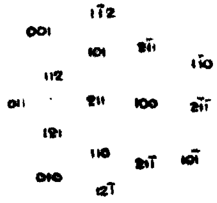
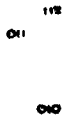


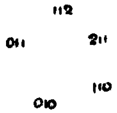
FIGURE 5. Barium on tungsten (thin deposit).



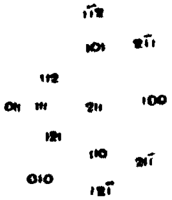
a/CLEAN TUNGSTEN
5870 V.



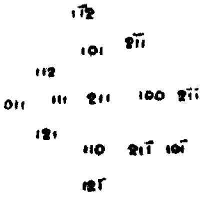
b/INITIAL DEPOSIT OF
BARIUM 2400 V.



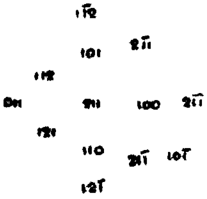
c/770°K - 30 SECONDS
MIGRATION PROCEEDS
2340 V.



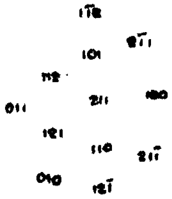
d/1 MINUTE AT 850°K
MIGRATION COMPLETE
2340 V.



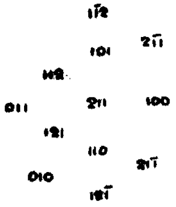
e/HEATING AT 900°K
REPOSITIONING OCCURS
2340 V.



f/AFTER HEATING AT 1000°K
FOR 5 MINUTES
EVAPORATION OCCURS. 3200V.



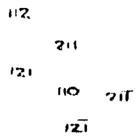
g/AFTER HEATING AT
1200°K. VERY SPARSE
COVERING OF BARIUM
4500 V.

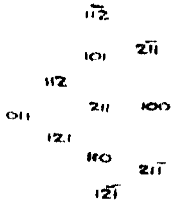


h/AT 1400°K PRACTICALLY
ALL THE BARIUM HAS
EVAPORATED 5400 V.

a) HEAVY DEPOSIT OF BARIUM.

b) ON HEATING AT 700 °K BARIUM MIGRATES, AND CRYSTALLITES FORM.





c) AT 850°K, SURFACE COMPLETELY COVERED IN ONE MINUTE AND CRYSTALLITES VERY EVIDENT.

d) AT 900 °K, CRYSTALLITES DISAPPEAR, APPEARANCE IS LIKE 5 d.

a) THIN DEPOSIT OF BARIUM 2700 V

b) NO MOVEMENT UNTIL HEATED TO 1000°K WHEN COALESCENCE AND WITHDRAWAL OF BARIUM EDGE APPEARED 2700V.

c) AT 1200 K FILM BREAKS UP INTO PATCHES 2700v

d) AT 1700°K CONSIDERABLE BARIUM HAS EVAPORATED, BUT NO MIGRATION HAS OCCURRED AT 1800°K, ALL BARIUM GOES, 3500V.

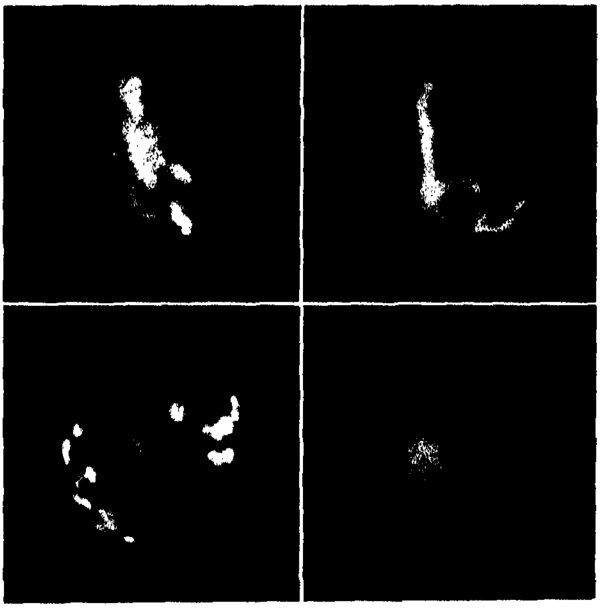


FIGURE 6. Barium on tungsten (thick deposit).

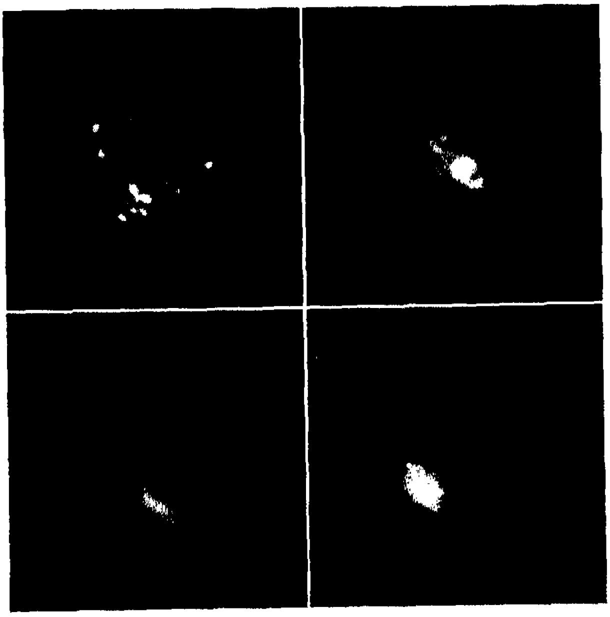


FIGURE 7. Non-migration of barium on tungsten.

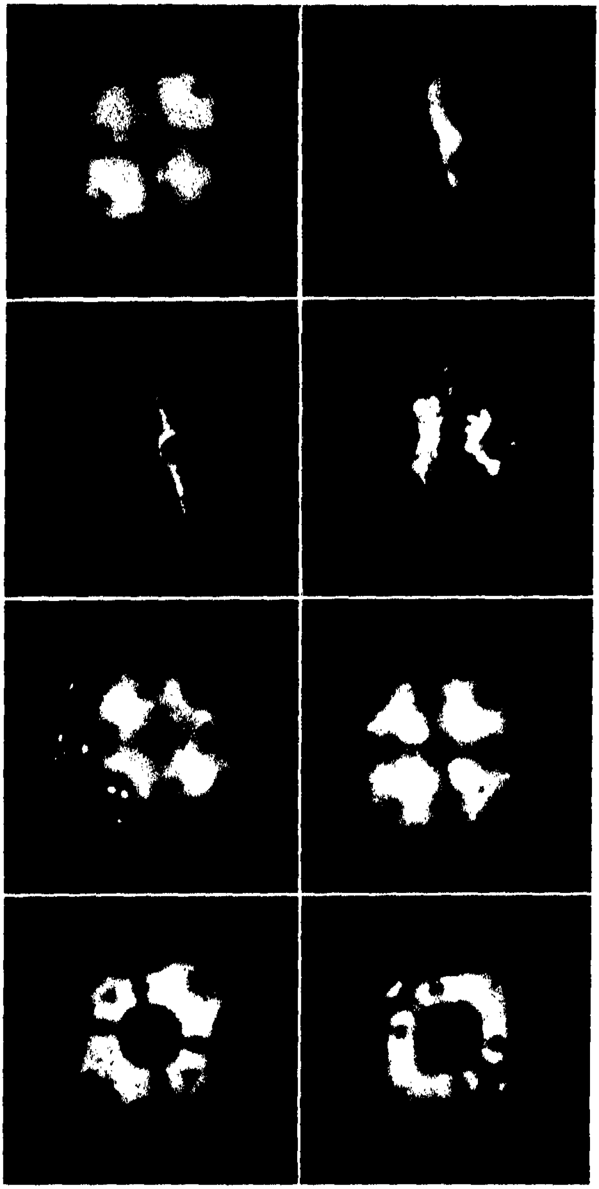
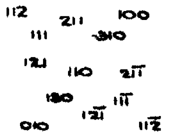


FIGURE 8. Barium on molybdenum

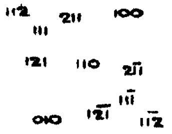


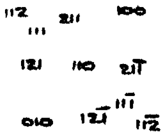
a) CLEAN MOLYBDENUM 10,000V.

b) INITIAL DEPOSIT OF BARIUM 3400 V.

c) FURTHER DEPOSITS OF BARIUM WITH MONATOMIC EDGE 3400V

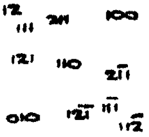
d) HEATING AT 600°K RESULTS IN MIGRATION AND CRYSTALLITE FORMATION 3400 V.

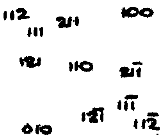




e) HEATING AT 600°K. RESULTS IN COMPLETE MIGRATION & BREAKDOWN OF CRYSTALLITES 3400V.

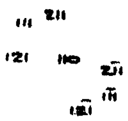
f) HEATING AT 650°K RESULTS IN UNIFORM COVERING & DISAPPEARANCE OF CRYSTALLITES 3400 V.



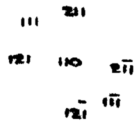


g) AT 900°K. EVAPORATION COMMENCES 4000 V.

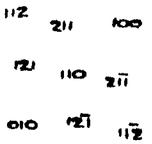
h) HEATING AT 950°K BARIUM LEAVES THE [110] REGION. 4000V.



i) HEATING AT 1200 °K
7 00 V.

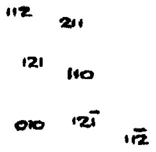


j) HEATING AT 1300 °K
SURFACE ALMOST
CLEAN. 7000 V

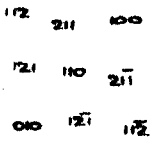


a) CLEAN BUILT-UP
MIGRATION.

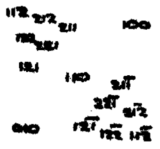
b) INITIAL DEPOSIT OF
BARIUM



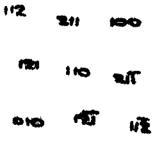
c) MIGRATION PROCEEDS
AT 700 °K.



d) MIGRATION COMPLETE
AFTER SEVERAL
MINUTES AT 700 °K.



e) EVAPORATION OCCURS
AT 850 °K.



f) HEATING AT 1300 °K
RESULTS IN ALMOST
COMPLETE REMOVAL
OF BARIUM AND ALL
RIDGES APPEAR.

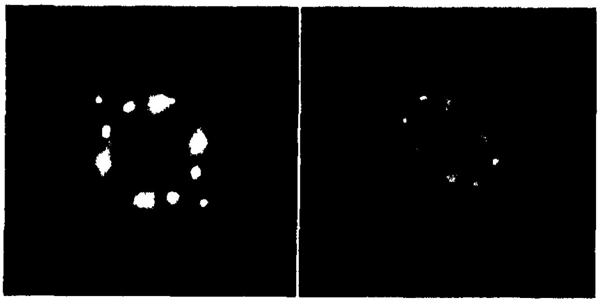


FIGURE 8

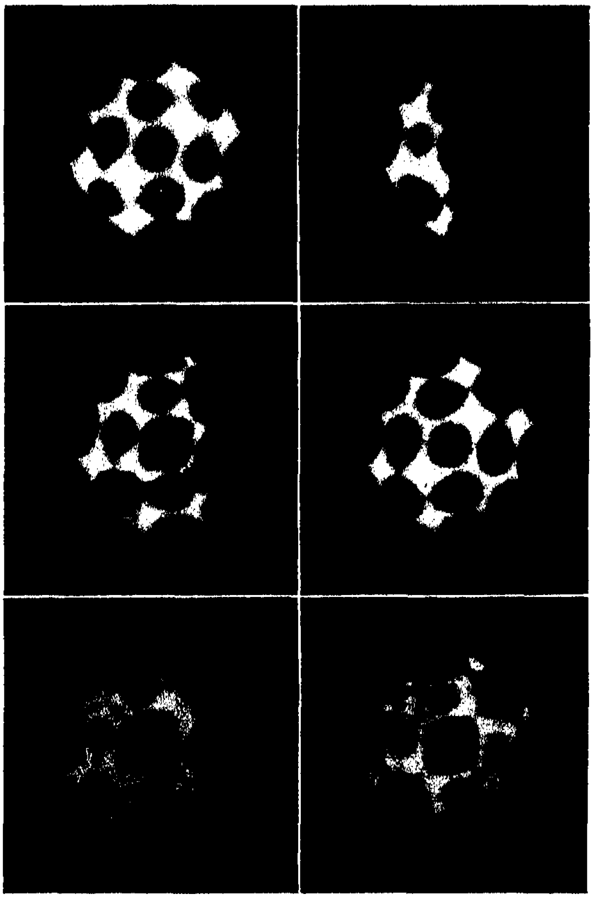


FIGURE 9. Barium on built-up molybdenum surface.

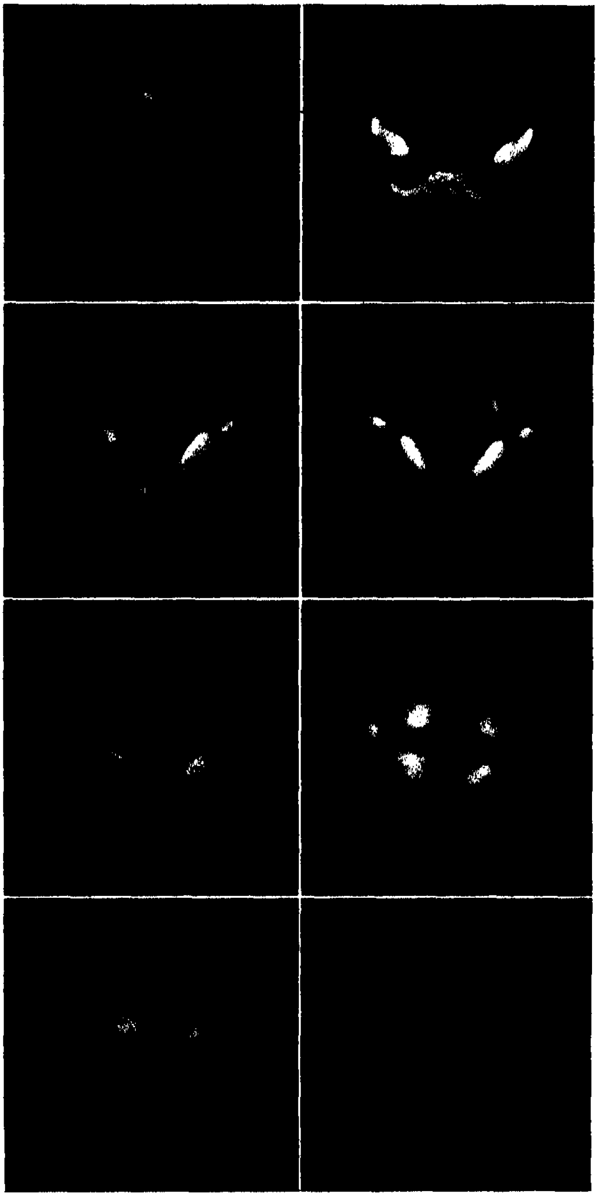
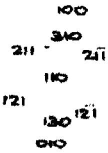


FIGURE 10. Sodium on molybdenum.

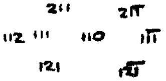


a) CLEAN SMOOTH MOLYBDENUM 12000 V.

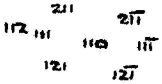
b) INITIAL DEPOSIT OF SODIUM 8000 V.

c) SOME MIGRATION AT 300 °K 8000 V.

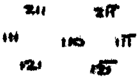
d) MIGRATION PROCEEDING AT 500 °K 6750 V.



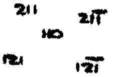
e) MIGRATION ABOUT COMPLETE AFTER 3 MINS. AT 500 °K 6750 V.



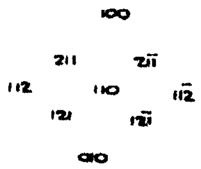
f) AT 600 °K, EVAPORATION 6750 V.



g) AT 700 °K EVAPORATION CONTINUES. 7000 V.



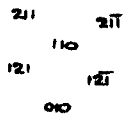
h) AT 850 °K, PRACTICALLY ALL THE SODIUM HAS EVAPORATED. 8000 V.



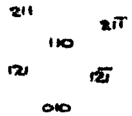
a) BUILT-UP CLEAN MOLYBDENUM SURFACE.

b). THICK Na DEPOSIT WITH MONATOMIC EDGE.

c) MIGRATION PROCEEDS AFTER 20 SECS. AT ROOM TEMPERATURE.



d) AFTER ONE MINUTE AT ROOM TEMPERATURE MIGRATION IS ALMOST COMPLETE,



e). AT 500°K MIGRATION COMPLETE & GIVING OUTLINE OF BUILT-UP PATTERN.

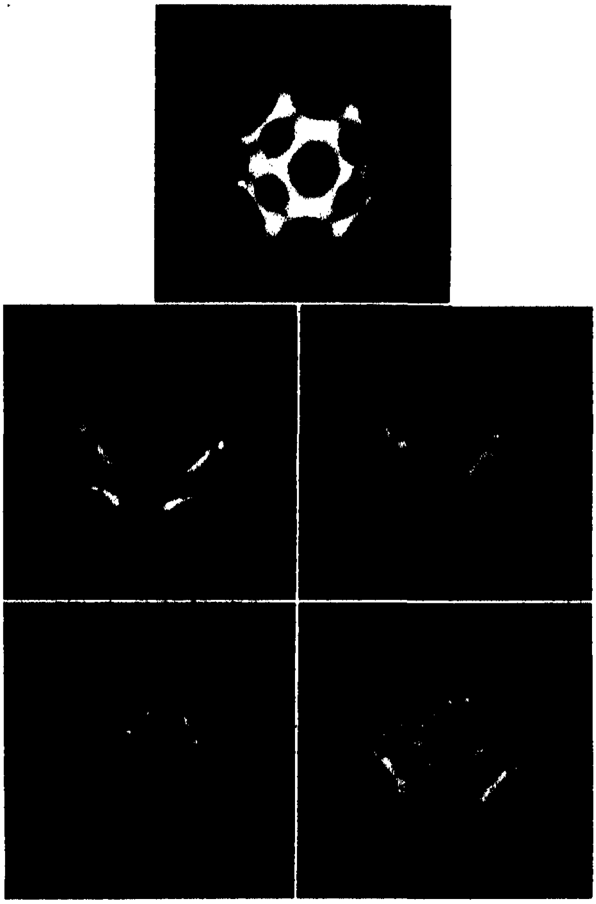


FIGURE 11. Thick sodium layer on built-up molybdenum.

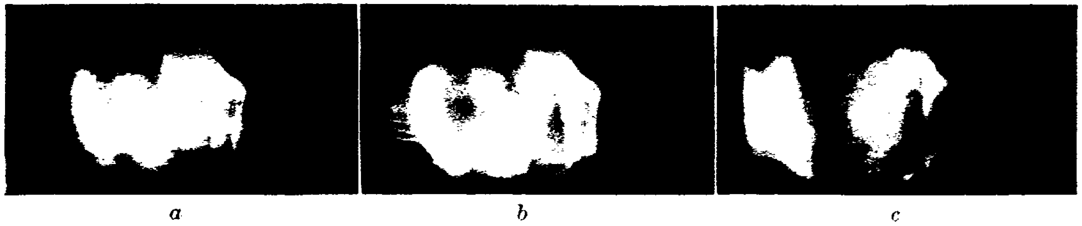


FIGURE 12. Migration of barium over single crystal tungsten wire.
a, initial deposit of barium. *b*, migration proceeds at 850° K. *c*, migration complete in 10 min. at 850° K

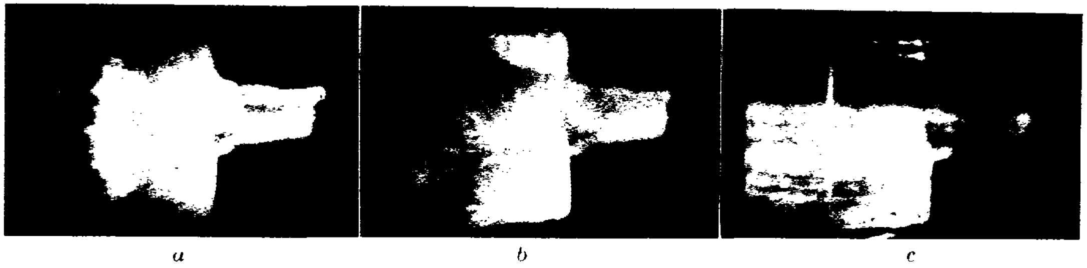


FIGURE 13. Migration of barium over polycrystalline tungsten.
a, initial deposit of barium. *b*, no movement observed until 1000° K, when migration occurred over some crystals accompanied by evaporation from others. *c*, migration complete on many crystals after 2 hr. at 1100° K.

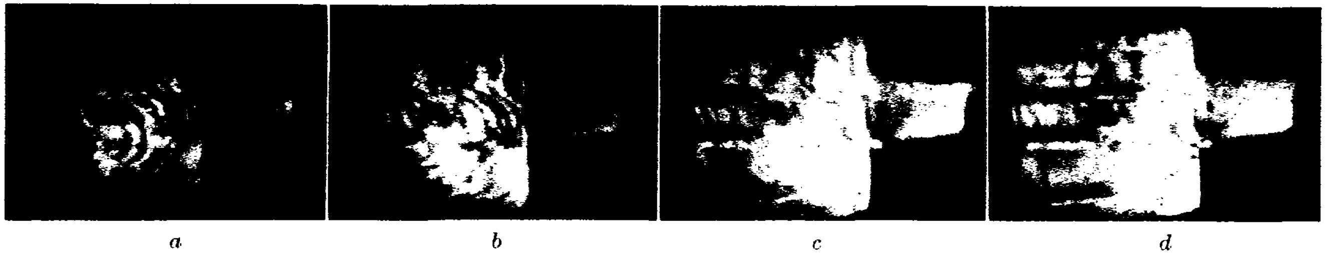


FIGURE 14. Migration of barium on a polycrystalline tungsten filament heated to obtain same grain growth.
a, initial deposit on large grain surface. *b*, movement apparent at 900° K. Appearance after 1½ hr. *c*, after ¾ hr. at 950° K, migration complete on some crystals. *d*, further heating at 1000° K gave no further migration, but some evaporation.

The detection and inhibition of free radical chain reactions

By J. R. E. Smith and C. N. Hinshelwood, F.R.S.

(*Received* 16 *September* 1941)

Part I. Comparison of nitric oxide and propylene as inhibitors

The reduction by propylene of the rate of pressure increase in the decomposition of propaldehyde at 550° has been shown by chemical analysis to represent a true inhibition of the reaction, and not to be due in an important degree to an induced polymerization of the propylene.

With propaldehyde and with diethyl ether the limiting values to which the decomposition rates are reduced by nitric oxide and by propylene respectively are the same, although much more propylene is required to produce a given degree of inhibition. From this it is concluded that the limiting rates are more probably those of independent non-chain processes, than those characteristic of stationary states where the inhibitor starts and stops chains with equal efficiency.

Since the decomposition of organic substances by free radical chain mechanisms was postulated (Rice & Herzfeld 1934), the detection of these chains has been a major problem. Staveley & Hinshelwood (1936*a*, *b*, *c*, 1937*a*, *b*) found that the rate of decomposition of some organic substances was considerably reduced by very small amounts of nitric oxide, which was assumed to break chains by combining with the free radicals. Increasing small amounts of nitric oxide reduce the rate to a limiting value which was believed to be the rate of an independent non-chain process. Still larger amounts of nitric oxide usually cause positive catalysis, which, however, is usually too small to affect the measurement of the maximum inhibition.

That nitric oxide can thus detect the presence of chains seems to be generally accepted. But the possibilities remain (1) that some radicals combine with nitric oxide too slowly for appreciable inhibition, and (2) that nitric oxide starts chains as well as stopping them. Rice & Polly (1938) assume that the maximally inhibited reaction is itself a modified chain mechanism in which the nitric oxide plays an essential part.

Thus two matters remain in some doubt: whether lack of inhibition is evidence of a non-chain reaction, and whether the maximally inhibited reaction is itself an internal rearrangement reaction or still a modified chain reaction.

Rice & Polly (1938) observed that propylene lowers the rate of pressure change during the decomposition of several organic substances, and supposed that it acted as an inhibitor by combining with active free radicals. The present series of papers studies this effect in detail for three typical reactions, namely, the thermal decompositions of diethyl ether, propaldehyde and acetaldehyde. The first two are known to be partially inhibited by nitric oxide (Staveley & Hinshelwood 1937 *b*), and a quantitative comparison of the behaviour of nitric oxide with that of propylene might be expected to throw light on the questions raised above. It was proposed first to determine whether the limiting rate reached with nitric oxide bears any definite relation to that reached with increasing amounts of propylene.

It was first necessary to show that the effect of propylene on the rate of pressure increase is a true reduction in the rate of decomposition, and not an illusory effect due to a polymerization of the propylene which superposes a pressure decrease on the normal increase. Direct chemical analysis was made by the method of Friedmann, Cotonio & Shaffer (1927) of the unchanged propaldehyde at various stages of the decomposition in presence of propylene. The results (table 1) show that even with considerable amounts present, a pressure increase of 1 mm. corresponds to the decomposition of nearly 1 mm. of the aldehyde, as is found with the pure substance (Hinshelwood & Thompson 1926), and as is found with acetaldehyde (see later). In view of this result, rates of reaction will be given as the reciprocal of the time for a pressure increase of 50 % of the initial pressure.

TABLE 1. RELATION OF PRESSURE INCREASE TO AMOUNT OF PROPALDEHYDE DECOMPOSED IN PRESENCE OF 100 MM. PROPYLENE. INITIAL PRESSURE OF PROPALDEHYDE 200 MM.; TEMPERATURE 550° C

propaldehyde decomposed (by analysis) mm.	pressure increase mm.
41	41
65	63
75	81
112	118
135	141

The rate of decomposition is considerably reduced by the addition of propylene, both with ether (table 2, figure 1) and with propaldehyde (table 3, figure 2). As with nitric oxide, there is first a fall in rate to a well-defined limit which cannot be further reduced, and later, with large amounts of propylene, an increase. With ether this increase is too gradual to affect appreciably the estimated value of the limiting rate. With propaldehyde, however, the amount of propylene needed to cause a given degree of inhibition is greater, and the secondary effects caused by the large amounts of propylene somewhat obscure the limit, though for the higher initial pressures of the aldehyde the effect is not serious (figure 2).

The parallel experiments with nitric oxide, made under identical conditions, are recorded in tables 4 and 5. To minimize the effect of the secondary

TABLE 2. INHIBITION OF ETHER DECOMPOSITION BY PROPYLENE. INITIAL PRESSURE OF ETHER 100 MM.; TEMPERATURE 550° C

propylene pressure mm.	$1/t_{50} \times 10^3$ (sec.$^{-1}$)
0	6·53
5	4·35
10	3·60
20	2·88
50	2·25
100	2·02
200	2·17
300	2·28

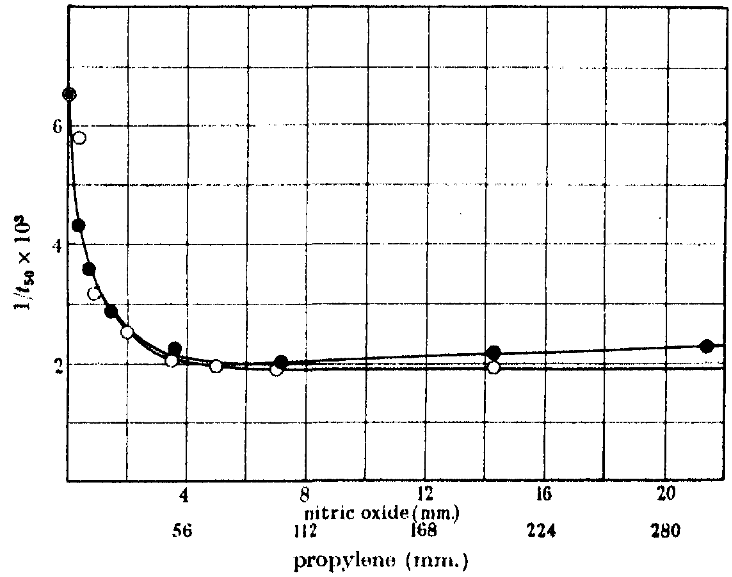


FIGURE 1. Inhibition of ether decomposition by nitric oxide (open circles): and by propylene (black circles).

TABLE 3. INHIBITION OF PROPALDEHYDE DECOMPOSITION BY PROPYLENE AT 550° C

propylene pressure mm.	$p_0 = 50$ mm. $1/t_{50} \times 10^3$	$p_0 = 160$ mm. $1/t_{50} \times 10^3$	$p_0 = 300$ mm. $1/t_{50} \times 10^3$
0	4·25	5·55	6·50
5	3·61	—	—
10	3·46	—	—
20	2·96	4·73	5·99
50	2·37	4·06	5·40
100	2·51	3·62	4·96
150	—	—	4·58
200	2·79	3·64	4·44
300	—	3·79	—

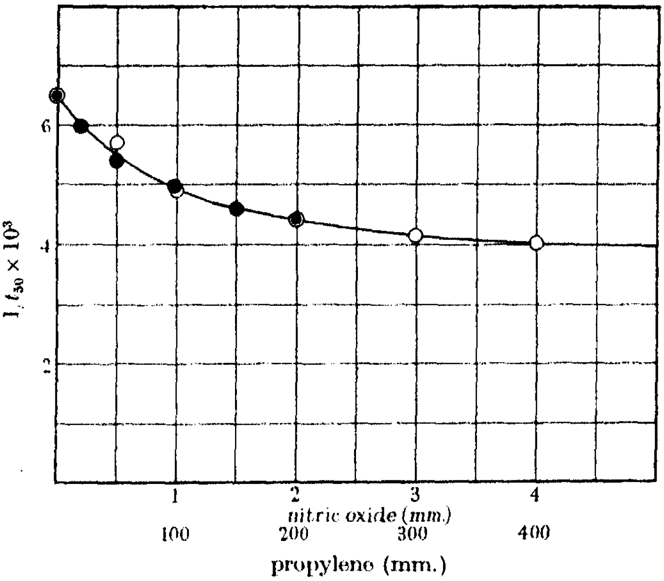


FIGURE 2. Inhibition of propaldehyde decomposition by nitric oxide (open circles): and by propylene (black circles).

TABLE 4. INHIBITION OF ETHER DECOMPOSITION BY NITRIC OXIDE. INITIAL ETHER PRESSURE 100 MM.; TEMPERATURE 550° C

nitric oxide pressure mm.	$1/t_{50} \times 10^3$ (sec.$^{-1}$)
0·00	6·54
0·37	5·80
0·90	3·17
2·00	2·53
3·50	2·05
4·97	1·94
7·1	1·91
14·3	1·92

TABLE 5. INHIBITION OF PROPALDEHYDE DECOMPOSITION BY NITRIC OXIDE. INITIAL PRESSURE OF ALDEHYDE 300 MM.; TEMPERATURE 550° C

nitric oxide pressure mm.	$1/t_{50} \times 10^3$ (sec.$^{-1}$)
0·0	6·50
0·5	5·70
1·0	4·89
2·0	4·43
3·0	4·16
4·0	4·01
6·0	3·96

pressure increase a large initial pressure (300 mm.) of propaldehyde was used.

Propylene is much less efficient as an inhibitor than nitric oxide. To compare the shapes of the inhibition curves, therefore, the pressure scales have been adjusted in the figures.

The results, as shown in the diagrams, reveal that propylene and nitric oxide appear to reduce the rates of decomposition to the same limiting values. This important experimental fact shows that the limiting rate is in all probability not an arbitrary value depending on the relative ability of the inhibitor to start and stop chains, but is the rate of an independent residual reaction in which the inhibitor itself does not play an essential part.

An alternative explanation would necessitate three assumptions: (1) that propylene could start chains in the manner assumed by Rice & Polly for nitric oxide; (2) that the relative ability of propylene to start and stop chains is almost identical with that of nitric oxide; (3) that this coincidence occurs both with ether and with propaldehyde. These assumptions would be rather arbitrary.

This conclusion is substantiated by the further result that the decomposition of propaldehyde maximally inhibited by nitric oxide cannot be further reduced in rate by the addition of propylene (table 6), as might be expected if the maximally inhibited reaction were a modified chain reaction. The only effect observed is a gradual increase in rate comparable with the secondary increase observed in the inhibition by propylene alone.

TABLE 6. EFFECT OF PROPYLENE ON THE DECOMPOSITION OF PROPALDEHYDE ALREADY MAXIMALLY INHIBITED BY NITRIC OXIDE. PROPALDEHYDE 160 MM., NITRIC OXIDE 2 MM.; TEMPERATURE 550° C

propylene pressure mm.	$1/t_{50} \times 10^3$ (sec.$^{-1}$)
0	3·13
35	3·25
100	3·46
150	3·60

Since the maximally inhibited reaction seems not to be a modified chain reaction characteristic of the particular inhibitor, the only remaining possibilities seem to be: (1) a non-chain internal rearrangement process, (2) a different chain process whose radicals are not absorbed either by nitric oxide or by propylene. Of these the first seems the more likely.

The behaviour of acetaldehyde with nitric oxide seemed at first to be in conflict with this conclusion, but the further investigations described in part IV have largely removed the discrepancy.

PART II. INHIBITION CURVES FOR THE THERMAL DECOMPOSITION OF DIETHYL ETHER IN PRESENCE OF PROPYLENE

The curves showing the inhibition of the ether decomposition as a function of nitric oxide concentration are identical for all initial pressures of ether. With propylene as inhibitor the corresponding curves are widely separated. The contrast can be explained by assuming a difference in the relative ease with which two types of radical combine with nitric oxide and propylene respectively.

That portion of the decomposition of diethyl ether which is subject to inhibition by nitric oxide has two important characteristics: (*a*) it is nearly of the first order; and (*b*) if ρ is the rate with a given amount of inhibitor, ρ_0 that with none and ρ_∞ the limiting rate, then $(\rho-\rho_\infty)/(\rho_0-\rho_\infty)$ plotted against [NO], the 'inhibition curve', is the same for all initial pressures of ether. (*a*) is consistent with a chain mechanism of the general type exemplified by equations (1)–(4) (Rice & Herzfeld 1934):

$$\begin{aligned}
&(1)\quad C_2H_5OC_2H_5 &&= CH_3 + CH_2OC_2H_5,\\
&(2)\quad CH_2OC_2H_5 &&= CH_3 + CH_3CHO,\\
&(3)\quad CH_3 + C_2H_5OC_2H_5 &&= C_2H_6 + CH_2OC_2H_5,\\
&(4)\quad CH_2OC_2H_5 + CH_3 &&= C_2H_5OC_2H_5,\\
&(5)\quad CH_2OC_2H_5 + NO &&= \text{inactive products.}
\end{aligned}$$

(*b*) requires that the chains are chiefly broken by a reaction of type 5, the radical which is chiefly attacked by the nitric oxide being that which normally undergoes the unimolecular decomposition in reaction 2 (Hobbs 1938).

TABLE 7. INHIBITION CURVES FOR PROPYLENE AND DIETHYL ETHER AT 550° C

pressure of propylene mm.	$(\rho-\rho_\infty)/(\rho_0-\rho_\infty)$		
	$p_0 = 35$ mm.	$p_0 = 100$ mm.	$p_0 = 300$ mm.
0	1·00	1·00	1·00
3	0·52	—	—
5	0·33	0·54	0·74
10	0·25	0·37	0·61
16	0·13	—	—
20	—	0·22	0·38
25	0·08	—	—
36	0·05	—	—
50	—	0·09	0·16
100	—	0·05	0·10

The inhibition curves obtained with propylene as inhibitor are given in figure 3 (data in table 7). In contrast with the nitric oxide curves, the propylene inhibition curves depend markedly on the initial pressure of the ether. This result can be explained by assuming that propylene reacts with both types of free radical present, namely, the CH_3 type and the $CH_2OC_2H_5$ type. This follows from the ordinary steady state calculations (see Hinshelwood 1940, pp. 89, 127). If the inhibitor is assumed to combine only with the $CH_2OC_2H_5$ type radical, then $(p-p_\infty)/(p_0-p_\infty)$ is a function of the concentration of the inhibitor but not of the initial pressure of the ether: if the inhibitor combines with the CH_3 type radical, then this expression is a function of the concentration of the inhibitor divided by that of the ether. The results with propylene are intermediate between these two ideal cases. Combination of propylene with both types of free radical may therefore be assumed to contribute to the chain breaking.

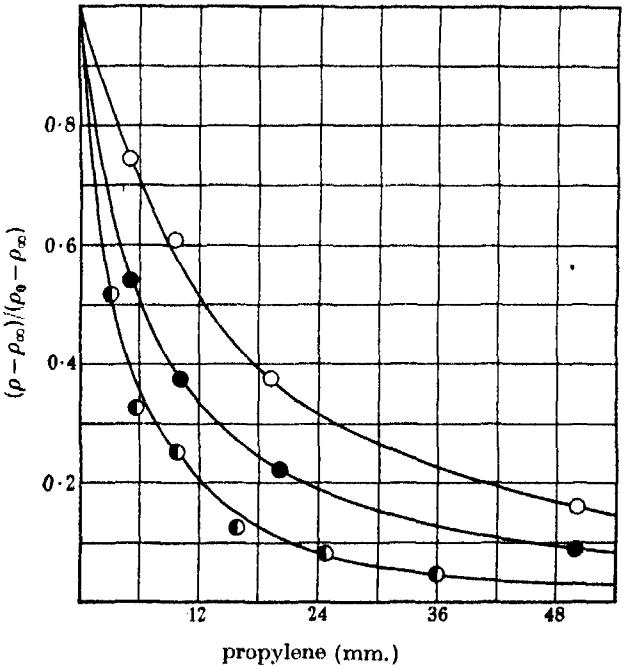


FIGURE 3. Inhibition curves for diethyl ether with propylene. (Open circles, 300 mm.; black circles, 100 mm.; half-shaded circles, 35 mm.)

As to the cause of this contrast, it is reasonable to suppose that with nitric oxide the chains end mainly by its combination with the larger radicals, reaction 2 being sluggish compared with reaction 3, so that a larger concentration of the heavier radicals is exposed to the attack. Propylene may well be less ready to combine with these large radicals, and consequently exerts relatively more of its effect on the lighter ones, although these are present in small concentration—greater concentrations of propylene being in fact required to produce the same degree of inhibition as nitric oxide.

PART III. INHIBITION CURVES FOR THE THERMAL DECOMPOSITION OF PROPALDEHYDE IN PRESENCE OF PROPYLENE

The thermal decomposition of propaldehyde is inhibited by propylene. The inhibited reaction is of an order between the first and the second. The 'inhibition curves', like those for the ether-propylene system and unlike those for the ether-nitric oxide system, depend upon the initial pressure of the propaldehyde.

The inhibition by nitric oxide and by propylene show that the decomposition of propaldehyde is at least partly a chain reaction. Detailed experimental results concerning the mechanism are lacking. Measurements are now available on the inhibition of the reaction by propylene.

TABLE 8. ESTIMATED VALUES FOR THE LIMITING RATES IN THE INHIBITION BY PROPYLENE OF THE DECOMPOSITION OF PROPALDEHYDE AT 550° C

initial pressure of propaldehyde, mm. ...	25	50	80	102	131
$1/t_{50}$ (sec.$^{-1}$) ...	1·66	1·92	2·28	2·60	2·86
initial pressure of propaldehyde, mm. ...	160	201	245	300	
$1/t_{50}$ (sec.$^{-1}$) ...	2·92	3·26	3·63	3·80	

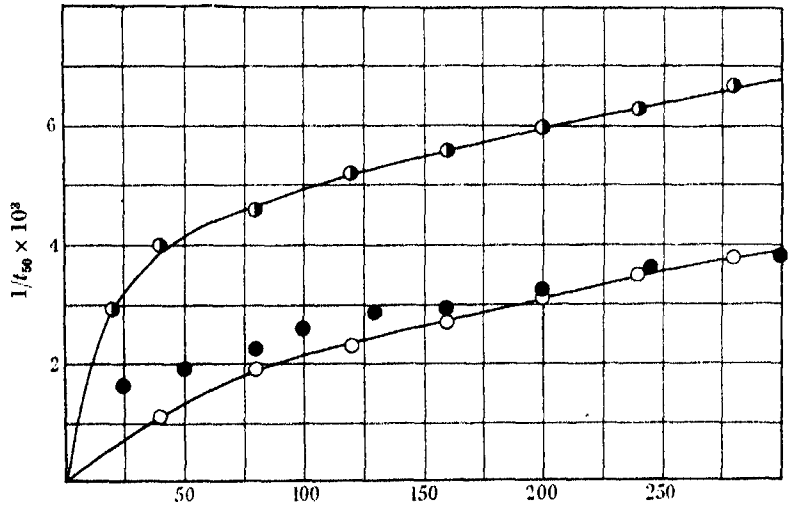


FIGURE 4. *Rates of normal propaldehyde decomposition and estimated limiting rates.* (Half-shaded circles, normal rate; open circles, inhibited by nitric oxide; black circles, inhibited by propylene.)

The determination of the limiting rate, ρ_∞, for low pressures of propaldehyde is made rather difficult by the secondary pressure increases occurring with large amounts of propylene. Nevertheless, it can be estimated with reasonable accuracy. Results for a series of initial pressures are given in table 8 and figure 4. This figure also includes (1) estimated values of the limiting rate of the reaction in presence of nitric oxide (from the results of Staveley & Hinshelwood 1936*c*), and (2) the rates of the uninhibited reaction (Winkler, Fletcher & Hinshelwood 1934). The apparent divergence of the two limiting rates at low initial aldehyde pressures may well be due to the errors of estimation. From figure 4 it can be seen (*a*) that the maximally inhibited reaction is of an order between the first and the second, and (*b*) that the inhibitable chain reaction is nearly of the first order (see also part V).

Figure 5 shows the separate 'inhibition curves' found for each of several initial pressures of propaldehyde.

TABLE 9. INHIBITION CURVES FOR PROPALDEHYDE WITH PROPYLENE AT 550° C

pressure of propylene mm.	$(p-p_\infty)/(p_0-p_\infty)$		
	$p_0 = 50$ mm.	$p_0 = 160$ mm.	$p_0 = 300$ mm.
0	1·00	1·00	1·00
5	0·73	—	—
10	0·66	—	—
20	0·47	0·69	0·82
35	0·29	—	—
50	0·20	0·44	0·62
100	—	0·27	0·46
150	—	—	0·34
210	—	—	0·27

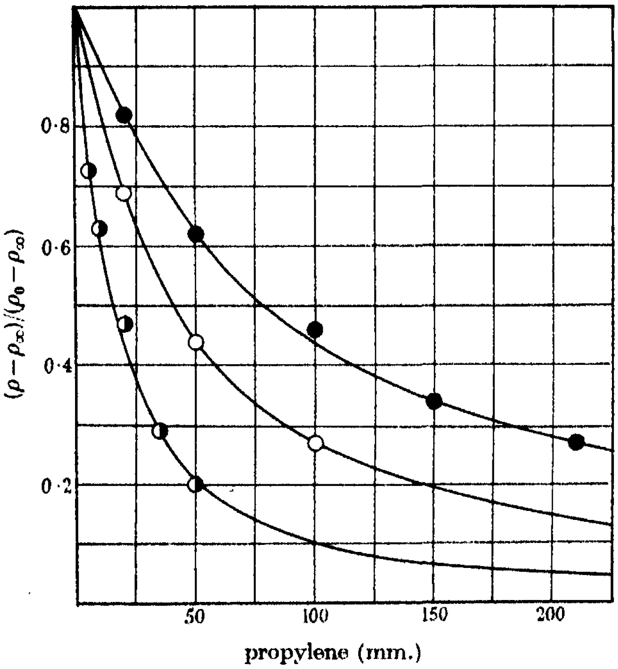


FIGURE 5. Inhibition curves for propaldehyde with propylene. (Half-shaded circles, 50 mm.; open circles, 160 mm.; black circles, 300 mm.)

The order of a chain reaction and the character of the inhibition curves depend upon the mechanism of chain ending (Hinshelwood 1940). Of the three common types, (1) bimolecular combination of like radicals, (2) bimolecular combination of two unlike radicals, as in the diethyl ether mechanism cited in part II, (3) termolecular combination of like radicals, the last is most consistent with the experimental facts for propaldehyde (compare part V).

PART IV. INFLUENCE OF PROPYLENE AND OF NITRIC OXIDE ON THE THERMAL DECOMPOSITION OF ACETALDEHYDE

The effect of propylene on the rate of pressure change in the thermal decomposition of acetaldehyde corresponds to a real inhibition of the reaction, as shown by direct chemical analysis. This inhibition therefore may be accepted as positive evidence for the existence of free radical chains in the decomposition. Re-examination of the negative evidence based upon the effect of nitric oxide indicates that in earlier measurements an inhibition may have been masked by a superimposed catalytic effect, and that the conflict in evidence yielded by nitric oxide and propylene is apparent only.

Previous evidence has clearly shown that the uninhibited decomposition of acetaldehyde is a *composite reaction*. The *reaction maximally inhibited* by propylene, probably an independent molecular reaction, has simpler kinetics: it is more nearly of an integral second order, the activation energy does not vary with pressure, and the rate corresponds to that of a collisional activation process with the activation energy in two square terms.

The inhibitable part of the reaction, taken to be the chain reaction, depends upon aldehyde and propylene pressure in such a way as to indicate that ternary and binary chain ending processes both occur. This explains the retarding effect of the reaction products which has hitherto been unaccounted for.

A theoretical chain mechanism for the decomposition of acetaldehyde was suggested by Rice and Herzfeld (1934) and has been supported by the following: (1) the approximate 3/2 order of the reaction (Letort 1937); (2) the analogy with the photochemical reaction and with the decomposition induced by free radicals from other sources (Allen & Sickman 1934); (3) theoretical arguments based upon spectroscopy which attribute more probability to a chain reaction than to an internal rearrangement (Burton & Davies 1939); (4) the detection of free radicals by the use of a radioactive indicator (Burton, Davies & Ricci 1940); (5) the reduction of the rate of pressure increase during decomposition by the addition of propylene (Rice & Polly 1938).

Fletcher & Hinshelwood (1933) found the order of the acetaldehyde decomposition to be variable, and postulated the existence of several simultaneously operating decomposition mechanisms (one or more of which, as is now clear, might have been chain reactions). This view was supported by the gradual variation in order in the series of aldehyde decompositions

(Hinshelwood, Fletcher, Verhoek & Winkler 1934) and by the variation of the activation energy with the initial pressure (Fletcher & Hinshelwood 1933; Smith 1939).

Against the chain theory were the arguments: (1) no inhibition by nitric oxide was observed by Staveley & Hinshelwood (1936*c*); (2) direct measurements of radical concentration by the para-hydrogen method indicated too few radicals for the required chains (Patat & Sachsse 1936); (3) no appreciable inhibition was produced by the presence of ethylene (Smith 1939); a similarity existed between the normal decomposition of acetaldehyde and the maximally inhibited decomposition of benzaldehyde, neither of which apparently suffered further inhibition by nitric oxide (Smith & Hinshelwood 1940).

Further evidence has now been sought by a more detailed study of the effect of propylene on the decomposition of acetaldehyde.

TABLE 10. EFFECT OF PROPYLENE ON RATE OF PRESSURE INCREASE IN DECOMPOSITION OF ACETALDEHYDE AT 550° C

propylene pressure mm.	$1/t_{50} \times 10^3$ (sec.$^{-1}$)			
	$p_0 = 50$ mm.	$p_0 = 100$ mm.	$p_0 = 200$ mm.	$p_0 = 400$ mm.
0	2·46	3·46	4·71	6·85
2	—	2·84	—	—
5	1·42	2·36	3·46	5·77
10	1·20	2·00	2·95	5·07
20	1·08	1·66	2·58	4·46
50	1·10	1·51	2·10	3·54
100	1·34	1·60	2·11	3·30
200	—	—	2·41	—

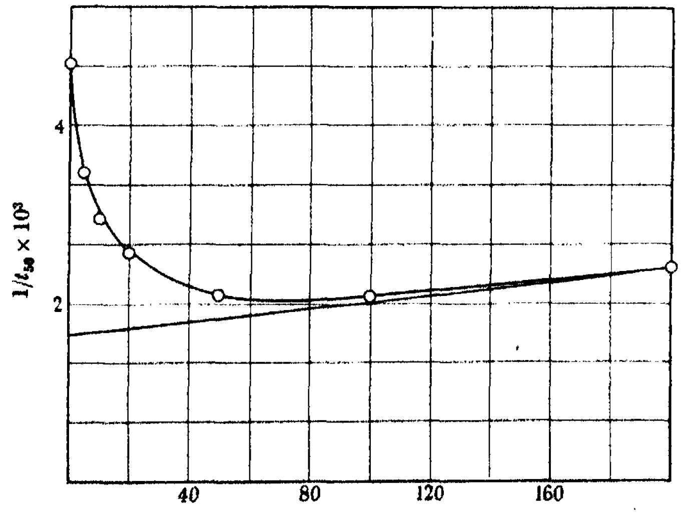


FIGURE 6. Inhibition of acetaldehyde decomposition by propylene.

The rate of pressure increase during decomposition is diminished by propylene (table 10, figure 6), and this is due to a real decrease in the reaction rate, as shown by direct chemical analysis (method of Friedmann, Cotonio & Shaffer 1927) (table 11). In presence of propylene 1 mm. pressure increase corresponds to a decomposition of 1 mm. aldehyde, as in the decomposition of the aldehyde alone (Winkler & Hinshelwood 1935). This constitutes positive evidence for a chain reaction, and appears to be in conflict with the lack of inhibition by nitric oxide observed by Staveley & Hinshelwood (1936*c*). A reinvestigation has therefore been made of the effect of nitric oxide (table 12, figure 7). At moderate initial pressures of aldehyde no appreciable inhibition is found, but at low initial pressures there is a small but appreciable inhibition. At higher pressures the first millimetre or two of nitric oxide causes little change in the rate: subsequent amounts, however, cause quite a marked increase. It seems likely that this is enough to mask a not inconsiderable degree of inhibition. Although the marked inhibition sought by the earlier observers does not occur, it appears probable

TABLE 11. RELATION OF PRESSURE INCREASE TO AMOUNT OF ALDEHYDE DECOMPOSED IN PRESENCE OF 100 MM. PROPYLENE. 200 MM. ALDEHYDE; 550° C

aldehyde decomposed by analysis mm.	pressure increase mm.
11	15
44	44
73	73
104	102
148	145

TABLE 12. EFFECT OF NITRIC OXIDE ON THE DECOMPOSITION OF ACETALDEHYDE

pressure of nitric oxide mm.	$p_0 = 50$ mm. p/p_0	$p_0 = 150$ mm. p/p_0	$p_0 = 400$ mm. p/p_0
0·00	1·00	1·00	1·00
0·16	—	0·97	—
0·21	—	—	0·98
0·31	0·96	0·98	—
0·50	0·89	0·99	0·98
0·80	0·90	—	—
1·00	—	1·04	1·03
1·20	0·93	—	—
2·00	1·01	1·24	1·30
2·76	—	—	1·45
3·00	1·16	1·49	—
5·00	1·70	—	—

that they underestimated the extent to which the positive catalysis might conceal it. If this is accepted, then the apparent conflict between the results with propylene and those with nitric oxide is removed. (Another explanation of the lack of appreciable inhibition by nitric oxide might have been that small amounts of it are very rapidly destroyed by the aldehyde. The rate of reaction of mixtures of ether and acetaldehyde is, however, much lowered by very small amounts of nitric oxide, showing that no such rapid removal occurs (table 13).)

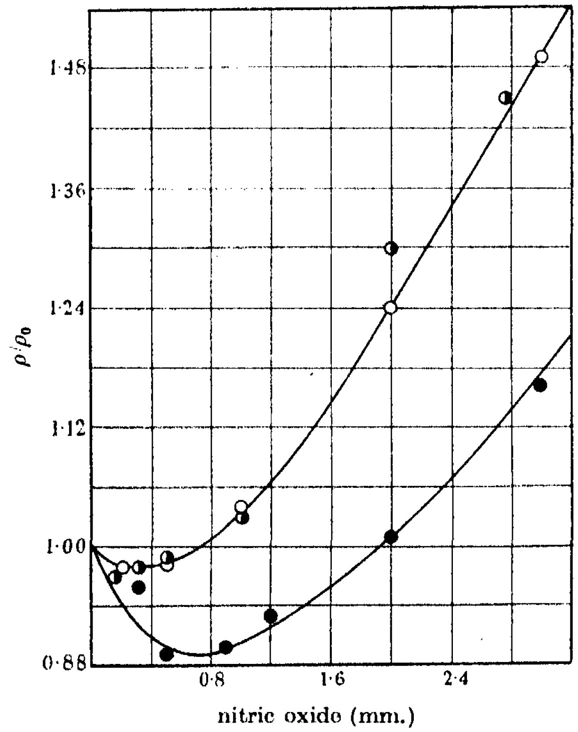


FIGURE 7. Effect of nitric oxide on acetaldehyde decomposition rate. (Black circles, 50 mm.; half-shaded circles, 150 mm.; open circles, 400 mm.)

TABLE 13. INHIBITION WITH ETHER (150 MM.)-ACETALDEHYDE (50 MM.) MIXTURES, 540° C

pressure of nitric oxide, mm. ...	0	0·15	0·30	0·72
$1/t_{50} \times 10^3$ (sec.$^{-1}$) ...	6·06	3·90	3·44	2·81

Repetition of previous experiments (Smith 1939), which showed inappreciable inhibition by ethylene, now show that under comparable conditions propylene is nearly 100 times more effective as an inhibitor than ethylene. (As to the probable reason for this compare Danby & Hinshelwood 1941.)

By analogy with ether and propaldehyde we may assume that the reaction maximally inhibited by propylene represents a non-chain decomposition of acetaldehyde: the properties of this reaction now become of great interest. Table 10 and figure 6 show that the limiting rates are well defined, although at the lower partial pressures of the aldehyde there is some inaccuracy introduced by the secondary pressure changes occurring with large amounts of propylene.

The limiting rates for a series of initial aldehyde pressures are given in table 14 and figure 8: the variation of the limiting rate, for various pressures, with temperature is given in table 15 together with the activation energies calculated by the method of least squares.

TABLE 14. ESTIMATED LIMITING RATES FOR DECOMPOSITION OF ACETALDEHYDE INHIBITED BY PROPYLENE: 550° C

initial pressure mm.	$1/t_{50} \times 10^3$ (sec.$^{-1}$)
50	0·70
100	1·12
152	1·53
200	1·66
248	2·00
305	2·14
350	2·70
400	2·84

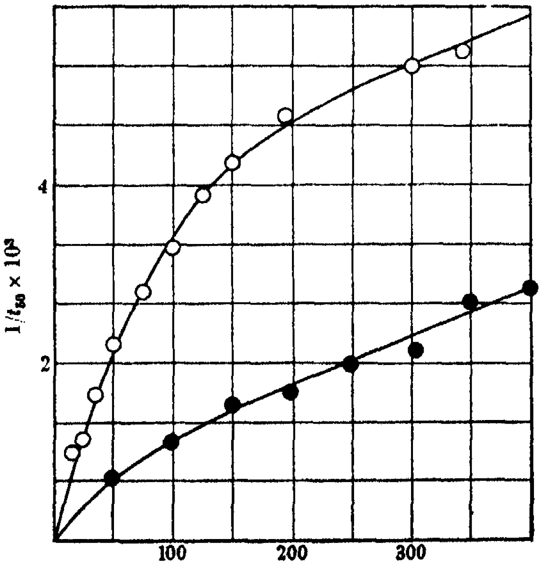


FIGURE 8. Normal and limiting rates of acetaldehyde decomposition. (Open circles, normal; black circles, limiting rate.)

The kinetics of the inhibited reaction are, as might be expected, rather simpler than those of the composite reaction. The order, as shown by the relation of rate and pressure in figure 8, is more nearly of the second order—and might even be precisely and integrally of this order if the inaccuracies attending the determination of the limiting rates could be entirely removed. The activation energy, unlike that of the total reaction, no longer varies with pressure. Its constancy is most clearly shown by the constancy with temperature change of the ratio of the rates at 75 and 350 mm. initial aldehyde pressure (table 15). With the activation energy given in the table the rate of the maximally inhibited reaction is approximately that of a collisional activation with the energy in two square terms (table 16). The properties of the reaction are considered further in a comparison of various aldehydes in part V.

TABLE 15. ACTIVATION ENERGIES FOR MAXIMALLY INHIBITED REACTION

initial pressure, mm.	activation energy, cal.
75	47,800
200	46,200
350	47,000

temp. °C	t_{50} (sec.) $p_0 = 75$ mm.	$p_0 = 200$ mm.	$p_0 = 350$ mm.	rate ratio 350 mm./75 mm.
505	—	—	2030	—
515	—	2346	—	—
520	3600	—	1216	2·96
535	2066	1108	687	3·01
550	1162	—	393	2·96
555	—	565	—	—
565	714	—	243	2·94
570	—	337	—	—
580	434	—	142	3·05
595	255	—	—	—

TABLE 16. COMPARISON OF ACTUAL RATE OF DECOMPOSITION OF ALDEHYDE WITH RATE OF COLLISIONAL ACTIVATION IN TWO SQUARE TERMS, FOR 100 MM. AND 550° C

	maximally inhibited reaction	'normal' reaction
$Ze^{-E/RT}$ (molecules per c.c. per sec.)	$1 \cdot 6 \times 10^{15}$	$1 \cdot 8 \times 10^{13}$
actual rate (molecules per c.c. per sec.)	$1 \cdot 3 \times 10^{15}$	$4 \cdot 1 \times 10^{15}$

The inhibitable portion of the reaction, the chain mechanism, can be seen from figure 8 to be of the first order at higher pressures and to change towards a higher order at lower pressures.

The inhibition curves with propylene are not of the ether-nitric oxide type, but vary with the initial aldehyde pressure (table 17, figure 9). These facts are most simply accounted for by the hypothesis that at the higher pressures the chains end by a termolecular process (first order reaction, inhibition curve a function of the aldehyde pressure), while at lower pressures there is appreciable probability of bimolecular chain ending (3/2 order reaction, inhibition curve still a function of the aldehyde pressure). This is further considered in connexion with other aldehydes in Part V. The contribution from termolecular chain ending processes would explain the slight inhibition of the reaction by its products (Letort 1937; Smith 1939), which increase the termolecular collisions. The effect was until now inadequately explained.

TABLE 17. INHIBITION CURVES FOR ACETALDEHYDE-PROPYLENE AT 550° C

pressure of propylene mm.	$p_0 = 50$ mm.	$p_0 = 100$ mm.	$p_0 = 200$ mm.	$p_0 = 400$ mm.
0	1·00	1·00	1·00	1·00
2	—	0·74	—	—
5	0·41	0·53	0·58	0·73
10	0·28	0·38	0·41	0·55
20	—	0·23	0·31	0·40
50	—	—	0·13	0·17
100	—	—	—	0·11

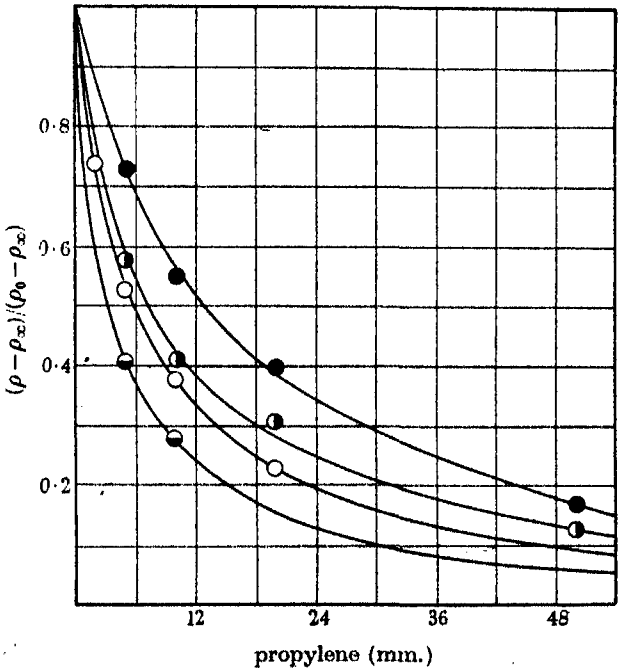


FIGURE 9. Inhibition curves with acetaldehyde and propylene. (Horizontally divided circles, 50 mm.; open circles, 100 mm.; vertically divided circles, 200 mm.; black circles, 400 mm.)

PART V. COMPARISON OF A SERIES OF ALDEHYDES

With the hypothesis that propylene and nitric oxide repress radical chains leaving an independent molecular reaction, a coherent picture can be formed of the decomposition mechanisms of simple aldehydes, which now show greater uniformity. The non-chain reactions seem to depend upon a bimolecular collisional activation, with superposition of more complex modes with the larger molecules. The chain mechanisms appear to be of two simple types distinguishable by a detailed study of the action of inhibitors.

The hypothesis that the decomposition of aldehydes, maximally inhibited by nitric oxide or by propylene, is a molecular process, while the inhibitable part of the reaction is an independent chain process leads, now that acetaldehyde can be brought into the scheme, to a picture of the kinetic relations which, if not final, is at least clearer and more coherent than before.

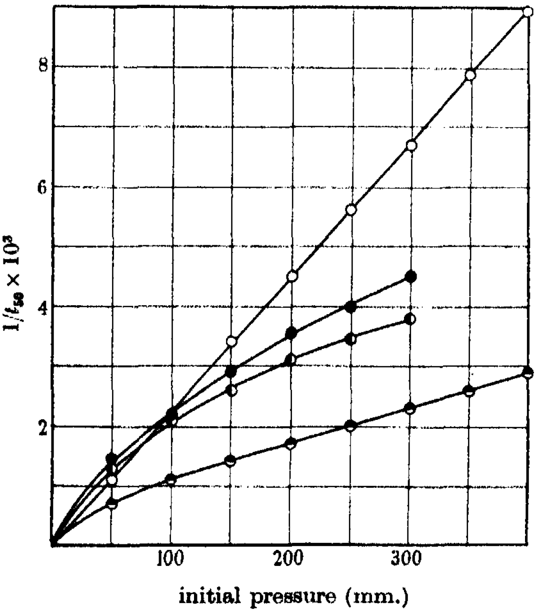


FIGURE 10. Rate of the normal formaldehyde decomposition and of the maximally inhibited decompositions of other aldehydes at 550° C. (Open circles, HCHO; horizontally divided circles, CH_3CHO; vertically divided circles, C_2H_5CHO; black circles, C_6H_5CHO.)

From the results of parts I–IV and of earlier papers the figures 10 and 11 have been constructed. Figure 10 gives the rates of the maximally inhibited reaction as a function of pressure for acetaldehyde, propionic aldehyde and benzaldehyde. The reactions are seen all to approximate to the second order, though by no means exactly. They do this, however, much more closely than the uninhibited reactions. The results for the normal decomposition of formaldehyde are included in the same diagram. This reaction appears to be integrally of the second order and free from chains. The greater regularity which now appears for the different aldehydes lends support to the initial hypothesis.

For formaldehyde and acetaldehyde the rates are of the same magnitude as the rate of a simple collision mechanism with the activation energy in two square terms: a mechanism which is consistent with the nearly second order nature of the decompositions. For benzaldehyde and propaldehyde the activation energies are higher, and the rates of reaction cannot be accounted for by activation in two square terms only. The operation of more complex activation mechanisms is perhaps revealed also in the deviation from the second order. The hypothesis of several modes of activation may still prove necessary to explain the results completely, even for the molecular part of the reaction. Complexity is also suggested by the variation with initial pressure of the activation energy for the maximally inhibited reaction of propaldehyde (Staveley & Hinshelwood 1936*c*) and of benzaldehyde (Smith & Hinshelwood 1940), an effect not observed with the maximally inhibited decomposition of acetaldehyde. For the series as a whole, however, there is now a closer approach to the original picture of a bimolecular collisional activation which Hinshelwood & Hutchison (1926) based upon the study of acetaldehyde at initial pressures where the contribution from the chain reaction is relatively less.

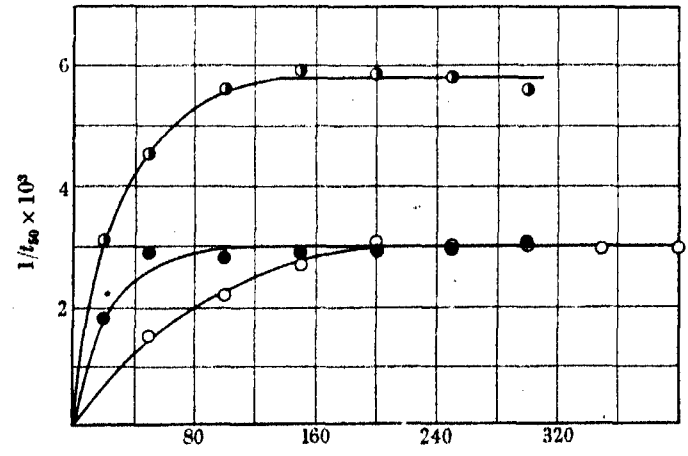


FIGURE 11. Differences between reciprocal half lives for normal and maximally inhibited decompositions of aldehydes, indicating the rates of the respective chain reactions. (Open circles, acetaldehyde; black circles, propaldehyde; divided circles, benzaldehyde.)

With regard to the kinetics of the chain processes, in figure 11 are plotted the differences between $1/t_{50}$ for the total reaction and for the maximally inhibited reaction. Although these are not strictly rates of the chain reaction, they are a close enough approximation to them for the diagram to leave little doubt that in all the reactions there is an approach to the first order. The fact that all the reactions yield the same result supports the hypothesis that the inhibitor does in fact isolate a separate reaction, and that the rate measured is not simply that of an arbitrarily modified chain reaction.

For the possible chain mechanisms there are two main types, already suggested for acetaldehyde and for benzaldehyde:

$$\begin{array}{lll}
\text{A.} & (1)\ CH_3CHO & = CH_3 + CHO, \\
 & (2)\ CH_3 + CH_3CHO & = CH_4 + CH_3 + CO, \\
 & (3)\ 2CH_3 & = C_2H_6, \\
 & (3a)\ 2CH_3 + X & = C_2H_6 + X. \\
\text{B.} & (1)\ C_6H_5CHO & = C_6H_5 + CHO, \\
 & (2)\ CHO & = CO + H, \\
 & (3)\ H + C_6H_5CHO & = C_6H_6 + CHO, \\
 & (4)\ CHO + H & = CO + H_2.
\end{array}$$

The essential differences from the kinetic point of view are that in B the spontaneous *decomposition of one of the participating radicals* is one of the rate determining steps, absent in A: and that in A there is the recombination of two *like* radicals, while in B there is recombination of two *unlike* radicals.

To distinguish between the two types there is the evidence (*a*) of the reaction order and (*b*) of the dependence of the 'inhibition curve' on the initial pressure of aldehyde.

A, with 3 as the chain ending step, gives a reaction of order 3/2: with 3*a*, the first order. It gives inhibition curves which depend upon the initial pressures. B gives the first order and inhibition curves independent of the initial pressure. With acetaldehyde and propaldehyde the inhibition curves are functions of the initial pressure and the order at higher pressures is nearly the first. This strongly points to type A, with termolecular collisions playing the main role in the chain breaking at higher pressures, a conclusion supported by the appreciable inhibition of the reaction by its products.

With benzaldehyde the inhibition curves coincide for different initial pressures and a mechanism of type B must be assumed to predominate. Probably each type of mechanism contributes in some degree with each aldehyde, the difference in reactivity of the alkyl radicals and of the phenyl radical being quite enough to explain why the preponderance is shifted on passing from the aliphatic to the aromatic series.

REFERENCES

Allen & Sickman 1934 *J. Amer. Chem. Soc.* **56**, 2031.
Burton & Davies 1939 *J. Chem. Phys.* **7**, 1075.
Burton, Davies & Ricci 1940 *J. Amer. Chem. Soc.* **62**, 265.

Danby & Hinshelwood 1941 *Proc. Roy. Soc.* A, **179**, 169.
Fletcher 1934 *Proc. Roy. Soc.* A, **146**, 357.
Fletcher & Hinshelwood 1933 *Proc. Roy. Soc.* A, **141**, 41.
Friedmann, Cotonio & Shaffer 1927 *J. Biol. Chem.* **73**, 342.
Hinshelwood 1940 *Kinetics of chemical change.* Oxford.
Hinshelwood, Fletcher, Verhoek & Winkler 1934 *Proc. Roy. Soc.* A, **146**, 327.
Hinshelwood & Hutchison 1926 *Proc. Roy. Soc.* A, **111**, 380.
Hinshelwood & Thompson 1926 *Proc. Roy. Soc.* A, **113**, 221.
Hobbs 1938 *Proc. Roy. Soc.* A, **167**, 456.
Letort 1937 *J. Chim. phys.* **34**, 355.
Patat & Sachsse 1936 *Z. phys. Chem.* B, **31**, 105.
Rice & Herzfeld 1934 *J. Amer. Chem. Soc.* **56**, 284.
Rice & Polly 1938 *J. Chem. Phys.* **6**, 273.
Smith 1939 *Trans. Faraday Soc.* **35**, 1328.
Smith & Hinshelwood 1940 *Proc. Roy. Soc.* A, **175**, 131.
Staveley & Hinshelwood 1936*a* *Nature, Lond.*, **137**, 29.
Staveley & Hinshelwood 1936*b* *Proc. Roy. Soc.* A, **154**, 335.
Staveley & Hinshelwood 1936*c* *J. Chem. Soc.* p. 812.
Staveley & Hinshelwood 1937*a* *Proc. Roy. Soc.* A, **159**, 192.
Staveley & Hinshelwood 1937*b* *J. Chem. Soc.* p. 1568.
Winkler, Fletcher & Hinshelwood 1934 *Proc. Roy. Soc.* A, **146**, 345.
Winkler & Hinshelwood 1935 *Proc. Roy. Soc.* A, **149**, 355.

On the significance tests for the introduction of new functions to represent measures

By Harold Jeffreys, F.R.S.

(*Received* 24 *October* 1941)

Significance tests for the introduction of one new function or two orthogonal new functions to represent a series of measures have previously been derived as approximate formulae. The exact formulae have now been reduced to single integrals involving only the number of observations and one function of the measured values, and their evaluation has been carried out. The approximate formulae are as accurate as we need when the number of observations exceeds the number of unknowns by about 6, but the extended calculation also gives the results when the difference is smaller. An approximate treatment is given for the case where the normal equations for two harmonics are not orthogonal in the coefficients and the weights unequal. This is the usual practical case. The solution for the symmetrical case, which has been done accurately, can be adapted to the general one by including two extra factors, of which one is calculated and the other can easily be found in any given problem.

1. Most of the significance tests that I have developed (1939) are reduced to reasonably compact approximate formulae, valid when the number of observations is large. The only problems where exact formulae in finite terms have been obtained are for the comparison of a sampling ratio with

a predicted chance (1939, p. 195) and for the comparison of two Poisson parameters (p. 211). The omission is not often important, since in practice the number of observations usually *is* large, but it is sometimes small enough for the approximations to become meaningless, and it is desirable to know how large the number must be before the approximations become useful. Determinations for smaller values are also needed sometimes.

The function K grades the evidence for or against the null hypothesis in respect of one suggested type of departure from it. $K = 1$ means that the test affords no new information, and the null hypothesis may be definitely rejected if $K = 0{\cdot}01$ or less. In some cases $1/K$ is reduced to a single integral of a known function, notably the tests for the consistency of two standard errors and for intraclass correlation (1939, p. 229), but these are functions of three variables and extensive tabulation would be necessary. Even in the two cases where there are answers in finite terms, K is a function of three variables, the two sample numbers and the suggested ratio. All tests involving χ^2 are essentially approximations, since Stirling's theorem or the method of steepest descents must be used before χ^2 appears at all.

The test reduces exactly to a function of two variables (of which the number of observations is one) for the introduction of new functions to fit a set of measures, when the standard error is initially unknown and there is no previous information about how much of the variation will be accounted for by the systematic effect sought. If only one new function is suggested, K is a function of the number of observations and of t, the ratio of the estimate of the coefficient to the usual estimate of its standard error. If there are more than one, K will still be a function of n and a generalized t^2, provided that the normal equations are orthogonal with regard to all the coefficients to be found and give the same weight to all.

2. We consider first the former problem. The exact solution is (Jeffreys 1939, p. 200, equations (11) and (12))

$$P(q \mid \theta H) \propto \int_0^\infty \sigma^{-n-1} \exp\left(-\frac{ns^2}{2\sigma^2}\right) d\sigma, \tag{1}$$

$$P(\sim q \mid \theta H) \propto \int_0^\infty \int_{-\sigma}^{\sigma} (\sigma^2-\alpha^2)^{-\frac{1}{2}n} \exp\left[-\frac{n\{s^2-a^2+(\alpha-a)^2\}}{2(\sigma^2-\alpha^2)}\right] \frac{d\sigma\, d\alpha}{2\sigma^2}, \tag{2}$$

where q is the null hypothesis, θ the data, H the statement of previous ignorance, and $\sim q$ the proposition that a new function f with coefficient α needs to be introduced. (A factor is understood to be included in the function so that Σf^2 over all observations $= n$.) a is the least squares estimate of α, s^2 the mean square deviation of the observed values from 0. The factor of proportionality is the same in both equations. We put

$$s/\sigma = k; \quad a/s = b; \quad \alpha/\sigma = u. \tag{3}$$

Then cancelling a factor s^{-n} we have

$$P(q \mid \theta H) \propto \int_0^\infty k^{n-1} \exp\left(-\tfrac{1}{2}nk^2\right) dk, \tag{4}$$

$$P(\sim q \mid \theta H) \propto \int_0^\infty \int_{-1}^1 \tfrac{1}{2}k^{n-1}(1-u^2)^{-\frac{1}{2}n} \exp\left[-\frac{n\{(k-bu)^2+(1-b^2)\,u^2\}}{2(1-u^2)}\right] dk\,du. \tag{5}$$

Evaluating the first of these we have

$$\begin{aligned} \frac{1}{K} &= \frac{P(\sim q \mid \theta H)}{P(q \mid \theta H)} \\ &= \frac{n^{\frac{1}{2}n}}{2^{\frac{1}{2}n}(\frac{1}{2}n-1)!} \int_0^\infty \int_{-1}^1 k^{n-1}(1-u^2)^{-\frac{1}{2}n} \\ &\qquad \times \exp\left[-\frac{n\{(k-bu)^2+(1-b^2)\,u^2\}}{2(1-u^2)}\right] dk\,du. \end{aligned} \tag{6}$$

Now in the notations of Airey, Irwin & Fisher (1931)

$$I_n(x) = \frac{1}{\sqrt{(2\pi)}} \int_0^\infty \frac{t^n}{n!} \exp\left(-\tfrac{1}{2}(t+x)^2\right) dt = \frac{1}{\sqrt{(2\pi)}} Hh_n(x), \tag{7}$$

$$Hh_{-1}(x) = \exp\left(-\tfrac{1}{2}x^2\right); \quad Hh_n(x) = \int_x^\infty Hh_{n-1}(x)\,dx \quad (n \geqslant 0). \tag{8}$$

These writers have tabulated $Hh_n(x)$ for n up to 11 and $x = -7{\cdot}0$ to $+7{\cdot}0$ at intervals of $0{\cdot}1$. Then

$$\frac{1}{K} = \frac{(n-1)!}{2^{\frac{1}{2}n}(\frac{1}{2}n-1)!} \int_{-1}^1 \exp\left\{-\frac{n(1-b^2)\,u^2}{2(1-u^2)}\right\} Hh_{n-1}\left(\frac{\sqrt{n}\,bu}{\sqrt{(1-u^2)}}\right) du \tag{9}$$

$$= \frac{(n-1)!}{2^{\frac{1}{2}n}(\frac{1}{2}n-1)!} \int_{-\infty}^\infty \exp\left\{-\tfrac{1}{2}n(1-b^2)\,v^2\right\} Hh_{n-1}(\sqrt{n}\,bv)\,(1+v^2)^{-\frac{3}{2}}\,dv. \tag{10}$$

Formula (9) was used for most of the numerical integration, but there were some complications. For b near 1 the maximum of the integrand is for u near -1, and the argument of the Hh function may lie beyond -7. In this range, however, the function is practically a polynomial. In fact for x large and negative $Hh_{-1}(x) \to 0$, $Hh_0(x) \to \sqrt{(2\pi)}$, and other functions can be derived by the recurrence formula (Fisher 1931, equation (13))

$$(n+1)\,Hh_{n+1}(x) + xHh_n(x) - Hh_{n-1}(x) = 0. \tag{11}$$

Some of the earlier functions, for $x < -7{\cdot}0$, are as follows, within the accuracy of the table:

$$\left.\begin{aligned} Hh_1(x) &= -x\sqrt{(2\pi)}, \\ Hh_2(x) &= \tfrac{1}{2}\sqrt{(2\pi)}\,(x^2+1), \\ Hh_3(x) &= -\sqrt{(2\pi)}\,(\tfrac{1}{6}x^3+\tfrac{1}{2}x), \\ Hh_4(x) &= \sqrt{(2\pi)}\,(\tfrac{1}{24}x^4+\tfrac{1}{4}x^2+\tfrac{1}{8}), \\ Hh_5(x) &= -\sqrt{(2\pi)}\,(\tfrac{1}{120}x^5+\tfrac{1}{12}x^3+\tfrac{1}{8}x). \end{aligned}\right\} \tag{12}$$

For $n = 3$ and 4, (9) was used up to a value of u large enough for $(v^2+1)^{-\frac{3}{2}}$ to be replaced by v^{-3}, and the contribution from higher values of v was found from (10) with this approximation. As has been explained already, it makes little difference to further procedure whether the odds are 3 or 4 to 1 against the null hypothesis, and it was considered sufficient to attempt an accuracy of 1 or 2 %.

K varies rapidly with b, but it was found that the differences of $\log(1-b^2)$ against $\log K$ permitted interpolation by divided differences to equal intervals of $\log K$. With the usual definition of t, we have

$$t^2 = \frac{(n-1)a^2}{s^2-a^2} = \frac{(n-1)b^2}{1-b^2}. \tag{13}$$

The integral (9) is finite for $n = 2$ even for $b = 1$. It is easy to verify that it is identically 1 for $n = 1$, and also for $n = 0$ if we use the limiting value of the first factor. This would be expected, and the case of $n = 1$ is already known. It means that we can get no evidence for or against the null hypothesis from a single observation, since its amount may be wholly random or wholly systematic and there is no way of knowing how much of it is random and how much systematic. Similarly we get none from no observations at all. These epistemological considerations provide a check on the analysis. But it was expected that for two observations sufficiently close together K could be very small; this is not so. The whole variation of K for $n = 2$ is only from 1·65 for $b = 0$ to 0·47 for $b = 1$. Hence we can never get strong evidence from two observations. The reason is connected with the fact that in the corresponding estimation problem the determination of the standard error from two or three observations is extremely bad, and leads to an infinite second moment for the probability law of the true value given the observations. Even for $n = 3$, though $1/K$ tends to infinity for $b = 1$, it does so logarithmically and extremely high values of t are needed before the evidence becomes strong. If two observations lie very close together, it is not strong evidence that σ is small, and therefore not strong evidence that the mean represents a systematic departure from 0.

Separate computations were done for each value of n up to 6 and 8, 10 and 12. The variation of $\log t$ with n for given K was such as to permit interpolation; the values of t corresponding to the same n and K were smaller than those given by the approximate formula, but the differences decreased nearly exponentially with n. The table could therefore be completed without difficulty. The results are as follows and replace and extend Table III of my book. Beyond $n = 20$ the table needs no change.

The maximum value of K is for $t = 0$; the approximate value is $(2n/\pi)^{\frac{1}{2}}$. The computed values are a little higher. If we write in this case

$$K = (2n'/\pi)^{\frac{1}{2}}; \quad n' = n+c$$

we find that for $n = 2$, $c = 2{\cdot}3$; $n = 6$, $c = 2{\cdot}7$; $n = 12$, $c = 2{\cdot}8$; $n = 20$, $c = 2{\cdot}8$. The correction is of little importance for higher values of n.

		t^2				
n	$K(t=0)$	$K=1$	$10^{-\frac{1}{2}}$	10^{-1}	$10^{-\frac{3}{2}}$	10^{-2}
2	1·65	2·13	—	—	—	—
3	1·86	2·22	30·5	5000	10^{10}	—
4	2·04	2·32	13·5	67·8	407	3230
5	2·19	2·40	10·0	29·6	85·3	241
6	2·34	2·46	8·4	20·4	44·9	94
7	2·48	2·50	7·7	16·3	31·8	59·3
8	2·61	2·55	7·2	14·2	26·2	43·0
9	2·73	2·60	6·9	13·3	22·0	34·7
10	2·84	2·67	6·7	12·4	19·9	30·1
11	2·95	2·75	6·6	11·8	18·6	27·4
12	3·07	2·8	6·5	11·4	17·6	25·2
13	3·17	2·8	6·5	11·1	16·9	23·7
14	3·27	2·8	6·5	10·8	16·3	22·5
15	3·37	2·9	6·4	10·6	15·8	21·7
16	3·46	2·9	6·4	10·4	15·2	20·9
17	3·55	3·0	6·4	10·3	14·9	20·2
18	3·64	3·0	6·3	10·1	14·7	19·7
19	3·73	3·0	6·3	10·1	14·3	19·2
20	3·81	3·1	6·3	10·0	14·0	18·8

It is interesting that the values of t^2 for $K = 1$ increase steadily with n, just as the corresponding values of χ^2 do. This of course is the level where the test is quite indecisive.

3. The analysis can also be completed for the testing of two orthogonal functions that enter together, as for a sine and cosine when the observations are at equal intervals of the argument from 0 to 2π. The exact forms are* (Jeffreys 1938, p. 174, equations (5), (6); 1939, p. 286, equations (18)–(21)):

$$P(q \mid \theta H) \propto \int_0^\infty \sigma^{-n-1} \exp\left(-\frac{ns^2}{2\sigma^2}\right) d\sigma, \tag{1}$$

$$P(\sim q \mid \theta H) \propto \int_0^\infty \iint \frac{1}{2\pi\gamma} \frac{d\sigma}{\sigma^2} \, d\alpha_1 \, d\alpha_2 \frac{1}{(\sigma^2-\gamma^2)^{\frac{1}{2}n}} \exp\left[-\frac{S(y-\alpha_1 f_1 - \alpha_2 f_2)^2}{2(\sigma^2-\gamma^2)}\right], \tag{2}$$

where $\alpha_1^2 + \alpha_2^2 = \gamma^2$ and the range of integration is over all values of α_1, α_2 such that $0 \leqslant \gamma \leqslant \sigma$: also

$$Sy^2 = ns^2, \tag{3}$$

$$S(y-\alpha_1 f_1 - \alpha_2 f_2)^2 = n(s^2 - 2\alpha_1 a_1 - 2\alpha_2 a_2 + \gamma^2), \tag{4}$$

$$na_1 = Syf_1; \quad na_2 = Syf_2, \tag{5}$$

$$\alpha_1 = \gamma\cos\phi, \quad \alpha_2 = \gamma\sin\phi, \quad a_1 = c\cos\lambda, \quad a_2 = c\sin\lambda. \tag{6}$$

---

* In the book the Greek and Roman letters are interchanged from the meanings in the paper to conform to the usual statistical practice. 'Degrees of freedom' in the title of the paper would also be replaced by 'new parameters' in the book, since 'degrees of freedom' in most statistical writings means the number of observations less the number of parameters evaluated.

Then
$$P(\sim q\,|\,\theta H) \propto \int_0^\infty \int_0^\sigma \int_0^{2\pi} \frac{1}{2\pi}\frac{d\sigma}{\sigma^2}\,d\gamma\,d\phi\,\frac{1}{(\sigma^2-\gamma^2)^{\frac{1}{2}n}} \times \exp\left[-\frac{n\{s^2-2\gamma c\cos(\phi-\lambda)+\gamma^2\}}{2(\sigma^2-\gamma^2)}\right]. \tag{7}$$

Put
$$s/\sigma = k; \quad \gamma/\sigma = u; \quad c/s = b, \tag{8}$$

and cancel a factor s^{-n}. Then

$$P(q\,|\,\theta H) \propto \int_0^\infty k^{n-1}\exp(-\tfrac{1}{2}nk^2)\,dk = \frac{2^{\frac{1}{2}n-1}}{n^{\frac{1}{2}n}}(\tfrac{1}{2}n-1)!, \tag{9}$$

$$P(\sim q\,|\,\theta H) \propto \int_0^\infty \int_0^1 \int_0^{2\pi} \frac{k^{n-1}}{2\pi}\frac{dk\,du\,d\phi}{(1-u^2)^{\frac{1}{2}n}} \times \exp\left[-\frac{1}{2}\frac{nk^2}{1-u^2}\left\{1-\frac{2bu}{k}\cos(\phi-\lambda)+\frac{u^2}{k^2}\right\}\right]. \tag{10}$$

Now
$$\int_0^{2\pi} e^{-t\cos(\phi-\lambda)}\,d\phi = 2\pi I_0(t) \tag{11}$$

($I_0(t)$ being here the Bessel function, not the function in (7)).

$$\frac{1}{K} = \frac{n^{\frac{1}{2}n}}{2^{\frac{1}{2}n-1}(\frac{1}{2}n-1)!}\int_0^\infty \int_0^1 k^{n-1}\frac{dk\,du}{(1-u^2)^{\frac{1}{2}n}}\exp\left[-\frac{1}{2}\frac{nk^2}{1-u^2}\left(1+\frac{u^2}{k^2}\right)\right] I_0\left(\frac{nkbu}{1-u^2}\right). \tag{12}$$

Also (Watson 1922, p. 394)

$$\int_0^\infty I_0(at)\exp(-p^2t^2)\,t^{\mu-1}\,dt = \sum_{m=0}^{\infty}\frac{(\frac{1}{2}a)^{2m}}{m!\,m!}\frac{(\frac{1}{2}\mu+m-1)!}{2p^{\mu+2m}}, \tag{13}$$

$$\frac{1}{K} = \frac{n^{\frac{1}{2}n}}{2^{\frac{1}{2}n}(\frac{1}{2}n-1)!}\int_0^1 \frac{du}{(1-u^2)^{\frac{1}{2}n}} \times \exp\left(-\frac{1}{2}\frac{nu^2}{1-u^2}\right)\Sigma\frac{(\frac{1}{2}n+m-1)!}{m!\,m!}\left(\frac{1}{2}\frac{n}{1-u^2}\right)^{-\frac{1}{2}n+m}(bu)^{2m}. \tag{14}$$

The series is a confluent hypergeometric function defined by

$$M(\alpha,\gamma,x) = 1+\frac{\alpha}{\gamma}x+\frac{\alpha(\alpha+1)}{\gamma(\gamma+1)}\frac{x^2}{2!}+\dots. \tag{15}$$

There is an identity
$$M(\alpha,\gamma,x) = e^x M(\gamma-\alpha,\gamma,-x), \tag{16}$$
and hence

$$\frac{1}{K} = \int_0^1 \exp\left\{-\frac{1}{2}\frac{n(1-b^2)\,u^2}{1-u^2}\right\} M\left\{1-\tfrac{1}{2}n,\,1,\,-\frac{1}{2}\frac{nb^2u^2}{1-u^2}\right\}du. \tag{17}$$

The function $M(\alpha,1,x)$ for $\alpha = -4$ to 4 at interval $\frac{1}{2}$ and $x = 0$ to 8 is given in the *British Association Report* for 1927. Unfortunately, the argument in (17) is negative, and the applicable table is that for $M(\frac{1}{2}n,1,x)$; but this varies very rapidly for large x and the intervals given are too wide for interpolation. I therefore computed tables of about four-figure accuracy of

$$M(1-\tfrac{1}{2}n, 1, -x) = 1+\frac{(\frac{1}{2}n-1)x}{1!1!}+\frac{(\frac{1}{2}n-1)(\frac{1}{2}n-2)}{2!2!}x^2+\dots \qquad (18)$$

For even n this is a polynomial. For odd n I found values for x up to 8 by dividing the B.A. values by e^x; for $n=3$ and 5 and x over 8 I used the asymptotic formula

$$e^{-x}M(\alpha,\gamma,x) \sim \frac{(\gamma-1)!}{(\alpha-1)!}x^{\alpha-\gamma} \times\left\{1+\frac{(\gamma-\alpha)(1-\alpha)}{x}+\frac{(\gamma-\alpha)(\gamma-\alpha+1)(1-\alpha)(2-\alpha)}{2!\,x^2}+\dots\right\}, \qquad (19)$$

three terms of which were sufficient. For $n=7$ the function was derived up to $x=35$ by the recurrence formula

$$\alpha M(\alpha+1,\gamma,x) = (x+2\alpha-\gamma)M(\alpha,\gamma,x)+(\gamma-\alpha)M(\alpha-1,\gamma,x), \qquad (20)$$

and for $x>35$ by (19). Arguments up to some hundreds arose in the computations for K near 0·01. I am indebted to Dr A. J. Thompson for letting me see some computations of his own for comparison.

An alternative form to (17) is

$$\frac{1}{K} = \int_0^\infty e^{-v}M\left(1-\tfrac{1}{2}n, 1, -\frac{b^2v}{1-b^2}\right)\frac{n(1-b^2)\,dv}{(2v)^{\frac{1}{2}}\{2v+n(1-b^2)\}^{\frac{3}{2}}}. \qquad (21)$$

This was used to deal with large values of the argument for $n=4$ and 5 just as the corresponding transformation was used for $n=3$ and 4 with one new function. Computations were made for n up to 8 and for 10 and 12. If $t^2 = \dfrac{(n-2)b^2}{1-b^2}$, t^2 is the sum of the squares of the ratios of the estimated coefficients to the usual estimate of their standard error. For other values of n, t^2 was found by interpolating the logarithms of the departures from the values found from the asymptotic formula used in Table IV of my book. K for $t=0$ and $n \geqslant 3$ is the same as for the same n and one new parameter. For $n=2$, t is indeterminate and $K=1$ as above stated.

As for one new function, the values of t^2 for given n and K are always less than those given by the approximate formula. The difference, however, would never affect K by a factor of $10^{-\frac{1}{2}}$ or less for given t^2 and $n \geqslant 8$, and is much less within the range where I thought it worth while to publish approximate values. At $n=30$ and beyond the approximate values need no change.

The general remarks made (Jeffreys 1939, pp. 356–60) about the use of these tests are equally applicable to these extended and more accurate tables. It should be mentioned also that conditions required for the use of the formulae are that there should be no previous information about the standard error or about whether the new functions are required. When n is large the former condition is of minor importance, since the information available is not usually as much as would be provided by a few extra observations in a directly relevant experiment; but it will become important if there is such information and n is small. The second condition excludes the kind of problem that often arises in agricultural and engineering testing, where it seems to me that the existence of the suggested difference is not seriously in doubt and the question is whether the estimated value has the right sign. I should regard this as a problem of estimation and not of significance (Jeffreys 1940).

	t^2				
n	$K=1$	$10^{-\frac{1}{2}}$	10^{-1}	$10^{-\frac{3}{2}}$	10^{-2}
3	4·1	—	—	—	—
4	4·3	97·5	4320	10^{14}	—
5	4·4	27·8	162	990	6340
6	4·4	18·6	56·7	174	537
7	4·4	14·9	36·6	80	177
8	4·4	13·0	28·1	53·5	105
9	4·5	11·8	23·1	42·7	78
10	4·5	11·1	20·6	34·7	54·5
11	4·6	10·7	19·0	30·3	45·5
12	4·7	10·4	17·9	27·5	40·2
13	4·8	10·2	17·1	25·6	36·8
14	4·9	10·0	16·6	24·4	34·2
15	4·9	10·0	16·2	23·4	32·3
16	5·0	9·9	15·7	22·5	30·6
17	5·1	9·8	15·4	21·7	29·2
18	5·1	9·8	15·1	21·1	28·2
19	5·2	9·7	14·9	20·6	27·2
20	5·3	9·7	14·6	20·2	26·5

4. *Allowance for non-orthogonality in the fitting of harmonics.* The calculations just given cover the case where the normal equations are orthogonal and the determinations of the two coefficients in the corresponding estimation problem are of equal weight. This would be true in a well-designed experiment, but not in a series of astronomical observations seriously interrupted by cloud, or in a test for agreement of two earthquake epicentres, where the distribution of observing stations is usually very unsymmetrical. In the analysis previously given (Jeffreys 1938) the non-orthogonal case was considered for the introduction of several new functions at once, and it was found that if there are more than two, and the coefficients are badly determined, the test can be adapted by omitting one of them altogether and applying a correcting factor to K; this rejection will improve symmetry, for which the approximate solution was found. This method fails if there are two new functions, which is the commonest case after that of one. We therefore need a test for the introduction of two new functions

when the determinations of the coefficients are of decidedly unequal weight. The analysis can be simplified a little by dispensing with the lemma and using the approximate solution obtained by a comparable method for the symmetrical case as a standard of comparison. Since the accurate solution for this case is now available, the adaptation to the more complicated one can be made with much more accuracy than would previously have been possible.

A slight change in the statement of the problem makes it agree better with the conditions that usually arise. If the new functions are

$$f_1 = \sqrt{2}\cos x; \quad f_2 = \sqrt{2}\sin x, \tag{1}$$

and their coefficients on the alternative hypothesis are α_1, α_2, the distribution of the observed values of x alone will usually be irrelevant to the ratio α_1/α_2. Then if σ is the standard error on the null hypothesis q and σ' on the alternative hypothesis $\sim q$, the expectation of the sums of squares of the observed values is $n\sigma^2$ on the null hypothesis and on the alternative hypothesis is

$$n\sigma'^2 + S(\alpha_1 f_1 + \alpha_2 f_2)^2 = n\sigma'^2 + b_{11}\alpha_1^2 + 2b_{12}\alpha_1\alpha_2 + b_{22}\alpha_2^2, \tag{2}$$

where b_{11}, b_{12}, b_{22} are the coefficients in the normal equations for α_1, α_2. We shall have

$$b_{11} + b_{22} = 2n. \tag{3}$$

Then (2) must be taken as the definition of σ^2 on $\sim q$, and α_1, α_2 are restricted by

$$b_{11}\alpha_1^2 + 2b_{12}\alpha_1\alpha_2 + b_{22}\alpha_2^2 \leqslant n\sigma^2. \tag{4}$$

But the condition that the values of x available for observation give by themselves no information about the direction of (α_1, α_2) requires us still to take the prior probability distribution

$$P(d\alpha_1, d\alpha_2 \mid \sim q, \sigma H) = A\,d\gamma\,d\phi, \tag{5}$$

where
$$\alpha_1 = \gamma\cos\phi, \quad \alpha_2 = \gamma\sin\phi, \tag{6}$$

and A is chosen so as to make the integral over the values permitted by (4) equal to 1. With a change of axes we can replace (4) by

$$B_1^2\beta_1^2 + B_2^2\beta_2^2 \leqslant n\sigma^2, \tag{7}$$

and
$$\begin{aligned}\frac{1}{A} &= \sqrt{n}\,\sigma\int_0^{2\pi}\frac{d\phi}{(B_1^2\cos^2\phi + B_2^2\sin^2\phi)^{\frac{1}{2}}}\\ &= 4\sigma\sqrt{n}\int_0^{\frac{1}{2}\pi}\frac{d\phi}{[\{\frac{1}{2}(B_1+B_2)\}^2\cos^2\phi + B_1B_2\sin^2\phi]^{\frac{1}{2}}}\end{aligned}$$

by Landen's transformation

$$= \frac{8\sigma\sqrt{n}}{B_1+B_2}F\left(\frac{B_1-B_2}{B_1+B_2}\right) = C\sigma, \tag{8}$$

F denoting the first complete elliptic integral. Then we find by the previous method

$$P(q\,d\sigma \mid \theta H) \propto \sigma^{-n-1} \exp\left(-\frac{1}{2}\frac{ns^2}{\sigma^2}\right) d\sigma, \tag{9}$$

$$\begin{aligned} P(\sim q\,d\sigma\,d\beta_1\,d\beta_2 \mid \theta H) &\propto \sigma'^{-n} \exp\left(-\frac{1}{2}\frac{ns'^2}{\sigma'^2}\right) \\ &\quad\times \exp\left[-\frac{1}{2\sigma'^2}\{B_1^2(\beta_1-b_1)^2+B_2^2(\beta_2-b_2)^2\}\right]\frac{d\sigma\,d\beta_1\,d\beta_2}{C\sigma^2\gamma}, \end{aligned} \tag{10}$$

and

$$s'^2 = s^2-(\beta_1-b_1)^2-(\beta_2-b_2)^2, \tag{11}$$

where b_1 and b_2 are the least squares estimates of β_1 and β_2. The most important case is where γ is small compared with σ and large compared with $\sigma/\sqrt{n}$. We put $\sigma = 1/k$ in (9) and $\sigma' = 1/k$ in (10); then

$$P(q\,dk \mid \theta H) \propto k^{n-1} \exp\left(-\tfrac{1}{2}nk^2s^2\right) dk, \tag{12}$$

$$\begin{aligned} &P(\sim q\,dk\,d\beta_1\,d\beta_2 \mid \theta H) \\ &\quad\propto k^n \exp\left[-\tfrac{1}{2}k^2\{ns'^2+B_1^2(\beta_1-b_1)^2+B_2^2(\beta_2-b_2)^2\}\right]\frac{dk\,d\beta_1\,d\beta_2}{Ck^3\sigma^3\gamma}. \end{aligned} \tag{13}$$

Integration with respect to β_1 and β_2 gives as an approximation for large n

$$P(\sim q\,dk \mid \theta H) \propto k^{n-1} \exp\left(-\tfrac{1}{2}nk^2s'^2\right)\frac{2\pi}{C(B_1B_2)}\frac{dk}{k^4\sigma^3\gamma}, \tag{14}$$

and integration with regard to k gives

$$K = \frac{P(q \mid \theta H)}{P(\sim q \mid \theta H)} = \frac{C}{2\pi}(B_1B_2)\,k^4\sigma^3\gamma\left(\frac{s'}{s}\right)^n, \tag{15}$$

where the quantities with small indices must be replaced by their values at the maximum of the integrand in (13), namely,

$$k = 1/s'; \quad \gamma^2 = b_1^2+b_2^2; \quad \sigma^2 = s'^2+(B_1^2b_1^2+B_2^2b_2^2)/n = s^2. \tag{16}$$

Then

$$K = \frac{C}{2\pi}(B_1B_2)\frac{\gamma}{s'}\left(\frac{s'}{s}\right)^{n-3}. \tag{17}$$

The standard errors of b_1 and b_2 are usually taken to be

$$s(b_1) = \left\{\frac{n}{(n-2)}\right\}^{\frac{1}{2}}\frac{s'}{B_1}; \quad s(b_2) = \left\{\frac{n}{(n-2)}\right\}^{\frac{1}{2}}\frac{s'}{B_2}, \tag{18}$$

and our generalized t^2 is given by

$$t^2 = \frac{b_1^2}{s^2(b_1)}+\frac{b_2^2}{s^2(b_2)} = \frac{n-2}{ns'^2}(B_1^2b_1^2+B_2^2b_2^2). \tag{19}$$

Then

$$K = \frac{4\sqrt{(n)}\,B_1B_2}{\pi(B_1+B_2)}F\left(\frac{B_1-B_2}{B_1+B_2}\right)\left(\frac{n}{n-2}\frac{b_1^2+b_2^2}{B_1^2b_1^2+B_2^2b_2^2}\right)^{\frac{1}{2}}t\left(1+\frac{t^2}{n-2}\right)^{-\frac{1}{2}(n-3)}. \tag{20}$$

In the symmetrical case $B_1 = B_2 = \sqrt{n}$, $F(0) = \frac{1}{2}\pi$, and K reduces to

$$K_0 = \frac{n}{(n-2)^{\frac{1}{2}}} t\left(1+\frac{t^2}{n-2}\right)^{-\frac{1}{2}(n-3)}, \tag{21}$$

which differs from the approximate form obtained before by having $n/\sqrt{(n-2)}$ for $\sqrt{n}$ and a factor t for $\sqrt{(t^2-1)}$. Both changes tend to increase K_0, and we have seen that the previous approximation made it too large. This approximation is therefore a little less accurate, but the difference is not important in the range $K < 1$, and what we need is the correcting factor for given t. This can conveniently be written

$$MN = \frac{4B_1 B_2}{\pi(B_1+B_2)\{\frac{1}{2}(B_1^2+B_2^2)\}^{\frac{1}{2}}} F\left(\frac{B_1-B_2}{B_1+B_2}\right)\left(\frac{\frac{1}{2}(b_1^2+b_2^2)(B_1^2+B_2^2)}{B_1^2 b_1^2+B_2^2 b_2^2}\right)^{\frac{1}{2}}, \tag{22}$$

in which the first two factors M depend only on B_1/B_2 and therefore on the asymmetry of the distribution of x, and both M and N reduce to 1 in the symmetrical case. We have also

$$B_1^2+B_2^2 = b_{11}+b_{22} = B; \quad B_1^2 B_2^2 = b_{11}b_{22}-b_{12}^2 = \Delta; \tag{23}$$

$$b_1^2+b_2^2 = a_1^2+a_2^2; \quad B_1^2 b_1^2+B_2^2 b_2^2 = b_{11}a_1^2+2b_{12}a_1 a_2+b_{22}a_2^2. \tag{24}$$

Put

$$\Delta = \tfrac{1}{4}\lambda^2 B^2. \tag{25}$$

Then

$$M = \frac{2\sqrt{2}}{\pi}\frac{\lambda}{\sqrt{(1+\lambda)}} F\left\{\sqrt{\left(\frac{1-\lambda}{1+\lambda}\right)}\right\}, \tag{26}$$

$$N = \left\{\frac{\frac{1}{2}(b_{11}+b_{22})(a_1^2+a_2^2)}{b_{11}a_1^2+2b_{12}a_1 a_2+b_{22}a_2^2}\right\}^{\frac{1}{2}}. \tag{27}$$

For complete symmetry $\lambda = 1$; for complete indeterminacy of one unknown $\lambda = 0$. The following table gives M as a function of λ.

λ	$M(\lambda)$	λ	$M(\lambda)$
1·0	1·00	0·05	0·114
0·9	0·94	0·04	0·095
0·8	0·87	0·03	0·076
0·7	0·80	0·02	0·054
0·6	0·72	0·015	0·042
0·5	0·64	0·010	0·030
0·4	0·55	0·008	0·025
0·3	0·45	0·006	0·019
0·2	0·33	0·004	0·014
0·15	0·27	0·002	0·009
0·10	0·20	0·001	0·004

The very small values of M for small λ may suggest that extremely badly distributed observations favour significance of the estimates, but two cautions are needed. First, the approximations assume that both b_1 and b_2 and their standard errors are small compared with s', so that the limits for β_1 and β_2 in (13) can be replaced by $\pm\infty$. Thus we must have B_1 and B_2

much greater than 1. If $B_2^2 = 4$, $B_1^2 = 2n-4$, we have $\lambda^2 \doteqdot 8/n$. Thus for the following values of n, λ must exceed the values given if the approximation is to be used safely.

n	λ
20	0·6
100	0·28
1000	0·09

I have not thought it worth while to attempt an approximation when $\lambda^2 < 8/n$, since the estimate of one unknown would be uncertain by something comparable with the standard error of one observation. If the observations relevant to a pair of harmonic coefficients are so badly distributed as to leave the phase almost indeterminate, the observer will presumably seek for more observations in the missing directions.

Secondly, if B_2 is small compared with B_1, b_2 will in general be of order s'/B_2 and b_1 not many times s'/B_1. Then the factor N will be of order $\frac{1}{4}B_1^2/B_2^2$ and therefore large. The factor N therefore tends to counteract the smallness of M when λ is small.

There is a slight difference between the analysis given here and that of the previous paper; in the latter the standard error on the hypothesis $\sim q$ was taken to be given by

$$\sigma'^2 = \sigma^2 - \alpha_1^2 - \alpha_2^2.$$

Thus the maximum likelihood estimate of σ^2 on $\sim q$ would be $s'^2 + a_1^2 + a_2^2$. But on q it would be $s'^2 + \dfrac{1}{n}(b_{11}a_1^2 + 2b_{12}a_1a_2 + b_{22}a_2^2)$, which is not the same except when the observations are symmetrically distributed. In accordance with the principle that parameters relevant on both hypotheses must be defined in such a way that they will receive the same maximum likelihood estimate on both, the present way of stating the problem is to be preferred.

The most serious cause of error in the approximations used is in the neglect in the integration of the variation of k^4 near the maximum of the integrand; but this has been allowed for completely by the accurate solution for the symmetrical case, and errors due to its having different values in unsymmetrical cases will differ much less. The suggested method is therefore to take

$$K = MNK_0,$$

where K_0 is to be taken from the table for the symmetrical case as for

$$t^2 = \frac{n-2}{ns'^2}(b_{11}a_1^2 + 2b_{12}a_1a_2 + b_{22}a_2^2) = \left(\frac{a_1 + b_{12}a_2/b_{11}}{s(a_1 + b_{12}a_2/b_{11})}\right)^2 + \frac{a_2^2}{s^2(a_2)},$$

M from the table as for λ defined by

$$\lambda^2 = 4\frac{(b_{11}b_{22} - b_{12}^2)}{(b_{11} + b_{22})^2},$$

and the special calculation of N from (27) is simple.

For analytical purposes it has been convenient to include a factor $\sqrt{2}$ with $\cos x$ and $\sin x$; but if this is not done and a pair of harmonics are fitted in the usual way there is no change in the way of applying the test. a_1 and a_2 will be multiplied by $\sqrt{2}$, b_{11}, b_{12}, b_{22} by $\frac{1}{2}$, and it is easy to verify that all the changes cancel. Hence the result can be applied directly to the values arising when the formation and solution of the normal equations is carried out in the ordinary way.

REFERENCES

Airey, J. R., Irwin, J. O. & Fisher, R. A. 1931 *British Association Mathematical Tables*, **1**, Table XV.
Jeffreys, H. 1938 *Proc. Roy. Soc.* A, **165**, 161–198.
Jeffreys, H. 1939 *Theory of probability*. Oxford: Clarendon Press.
Jeffreys, H. 1940 *Ann. Eugen., Camb.*, **10**, 48–51.
Watson, G. N. 1922 *Theory of Bessel functions*. Camb. Univ. Press.

Imperfections in the structure of cobalt

I. Experimental work and proposed structure

BY OLIVE S. EDWARDS AND H. LIPSON
Cavendish Laboratory, Cambridge

(*Communicated by A. J. Bradley, F.R.S.—Received* 14 *August* 1941)

(Plate 12)

X-ray powder photographs of hexagonal cobalt have a remarkable mixture of sharp and diffuse lines. The sharp lines are those with $\frac{1}{3}(h-k)$ integral and those with l zero; all others are more or less diffuse. An explanation of this is produced, based on the possibility of faults occurring in the structure because there are two different ways in which one close-packed layer of atoms can be put on to another. It is shown theoretically that if such faults are repeated at regular intervals the reflexions with $\frac{1}{3}(h-k)$ integral are not affected, but the positions and intensities of the other lines are modified although they remain sharp. These results suggest that a structure with random faults would give broadened reflexions, but the broadening is such that it would not show for lines with l zero on powder photographs.

Evidence from oscillation photographs and from measurement of a powder photograph is adduced in support of this theory.

INTRODUCTION

It has been known for a long time that the structure of cobalt is anomalous. The structure of any particular specimen depends on its previous history, and is often a mixture of face-centred-cubic and close-packed-hexagonal

Edwards & Lipson *Proc. Roy. Soc., A, volume* 180, *plate* 12

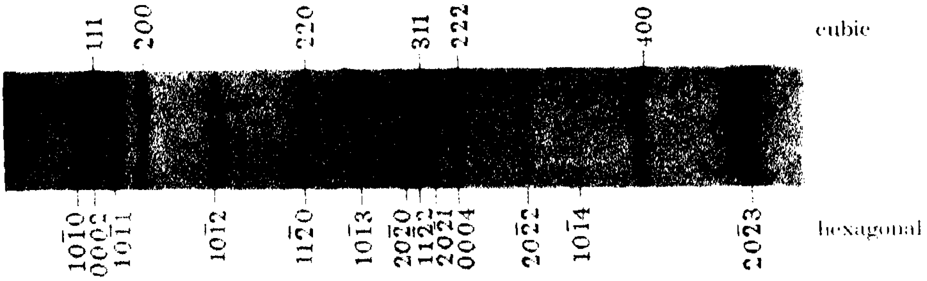


FIGURE 1

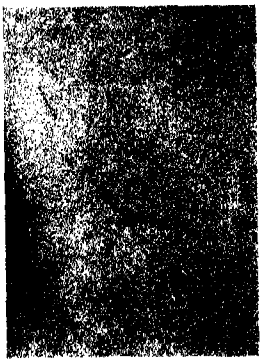


FIGURE 2*a*

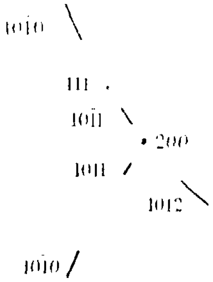


FIGURE 2*b*

(*Facing p.* 269)

forms (van Arkel 1939). This two-phase state has generally been ascribed to the presence of impurities (van Arkel 1939), but the authors have found the same tendency to be two-phase in some cobalt of purity 99·99 %.

During an investigation of this problem another interesting phenomenon was encountered, and a description of this, and a possible explanation of it, are the main subjects of this paper. The lines from the hexagonal cobalt are very often not all sharp (Hendricks, Jefferson & Shultz 1930) but vary in width. This variation is not simply a function of the angle of reflexion as it would be if it were due, for instance, to small particle size, nor can it be explained by particular shapes of small particles (von Laue 1936).

DESCRIPTION OF THE PHENOMENON

Powder photographs of cobalt annealed at temperatures below 400° C show a mixture of the patterns from the face-centred-cubic and hexagonal-close-packed structures. The lines of the former pattern are sharp, indicating that the crystal size is at least of the order of 10^{-5} cm. (Bragg 1933*a*); some of the hexagonal lines are also as sharp as this but others are very broad indeed. A typical photograph is shown in figure 1 (plate 12). There are difficulties in drawing general conclusions because of the overlapping of some of the hexagonal lines (those with $\frac{1}{3}(h-k)$ integral and l even) with some of the cubic lines, but it is probable that these lines are as sharp as the cubic lines. This is definitely so in the case of 0004 which is sufficiently far from its cubic companion 222 to be observed accurately. Lines with l zero are also sharp, and this would imply that the crystals can be thin only in the direction of the c-axis. This, however, is not in accordance with the other observations, particularly with the sharpness of 0004.

The broadening was also shown by reflexions from single crystals. These crystals were obtained in a rather unexpected manner. Some fine powder of the metal had been annealed at 1100° C for 11 days, and it was then found to consist of much larger grains. A Debye-Scherrer photograph showed two phases, but the lines were very 'spotty'; oscillation photographs were therefore taken of some individual grains, and these showed that the grains consisted of very few crystals. Two of the grains were examined in detail, and it was found that one of them contained two cubic crystals and four hexagonal crystals. The cubic crystals had relative orientations represented by twinning across a (111) plane; the four hexagonal crystals had their c-axes perpendicular to the $\{111\}$ planes of one of the cubic crystals, and their $[10\bar{1}0]$ directions were parallel to $[110]$ directions in the cubic crystal. These orientations are those which would be expected from the similarity of the two structures, and it has been shown by Nishiyama (1936) that the process of transformation is such that the cubic and hexagonal crystals are related in this way. The second specimen also had two cubic crystals, oriented as in the first specimen, but only two of the hexagonal crystals were present.

Part of one of the oscillation photographs is shown in figure 2 (plate 12). It will be seen that the hexagonal spots are broad, but now the rule that reflexions with l zero are sharp is not true, and it appears that the powder lines are sharp only because the reflexions lie along Debye-Scherrer lines. In addition, it can be seen that the directions of broadening of different spots are related. In figure 2 (plate 12) the impression is given that the spots are spread out along two lines, and since these lines run through the spots $10\bar{1}0$, $10\bar{1}1$, $10\bar{1}2$, they can be represented in reciprocal space as lines parallel to the c-axis. The complete representation of the reciprocal lattice must therefore be something like figure 3; certain reflexions are sharp, represented by spots in reciprocal space; others are strung out along lines parallel to the c-axis.

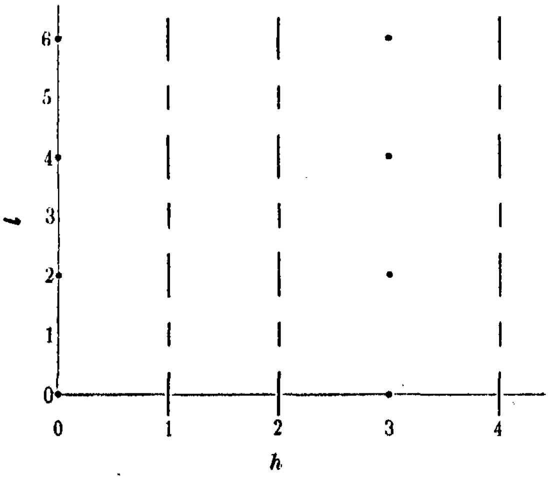


FIGURE 3. Section of the reciprocal lattice

THE PROPOSED STRUCTURE

The single crystal reflexions show that the underlying irregularities which produce the blurred lines are of a plate-like type perpendicular to the c-axis, but they must be such that they do not affect, for instance, the $000l$ reflexions. The close-packed structures can be considered as composed of a series of close-packed planes of atoms, each plane fitting on to the one below. There are three different types of planes (Bragg 1933*b*), and if the atomic positions are referred to hexagonal axes the representative positions as shown in figure 4 are:

A	000	or	$00\frac{1}{2}$
B	$\frac{1}{3}\frac{2}{3}0$	or	$\frac{1}{3}\frac{2}{3}\frac{1}{2}$
C	$\frac{2}{3}\frac{1}{3}0$	or	$\frac{2}{3}\frac{1}{3}\frac{1}{2}$

From spatial considerations only, each layer can fit on to either of the others. Thus there is an infinity of possible close-packed structures; the most important are the hexagonal, represented by the succession of planes *ABABAB*..., *BCBCBC*... or *CACACA*..., and the cubic, *ABCABCA*... or *CBACBAC*.... Since the relations between neighbouring atoms are the same in any close-packed structure, the differences between the structures must depend on the influence of the next nearest neighbours, and the fact that in cobalt the two structures can coexist suggests that here this influence is very small. We may thus imagine a structure in which occasional faults occur, a sequence *ABABAB*... changing to *BCBCBC*... and so on. The resulting regions of perfect crystallization would then be of the required plate-like character, but obviously reflexions like 000l would not be affected since the 0001 spacing is perfect over the whole crystal.

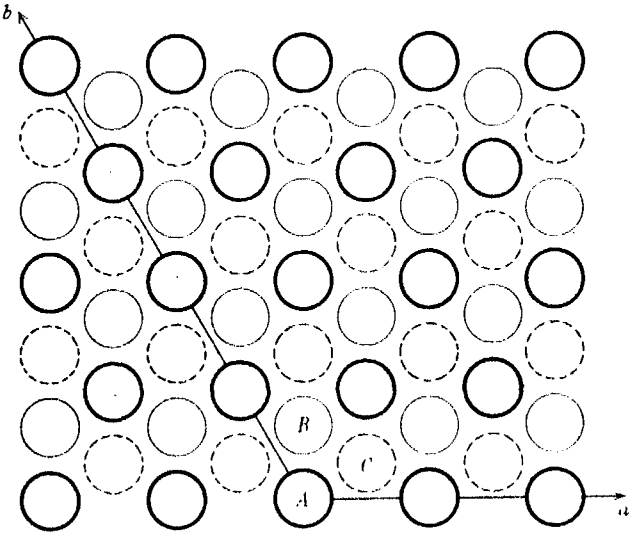


FIGURE 4. Relative atomic positions in the three different planes *A*, *B*, *C*.

In general, all reflexions involving the sum of the scattering factors of the atoms will be unaffected, but it is not clear what will happen to the others. The idea that they should be blurred resulted from consideration of the ordered alloy $AuCu_3$ (Sykes & Jones 1936) in a single crystal of which regions of order may exist which do not match (anti-phase domains). Those reflexions which are due to the sum of the scattering of the atoms—the main lines—are sharp; the others—the superlattice lines—are not. Jones & Sykes (1938) have attempted to deduce the size of the small regions of perfect order by treating the broadening as simply due to their small size, but there is no obvious justification for this.

MATHEMATICAL TREATMENT

The proposed structure cannot be treated by ordinary crystallographic methods since it is not periodic, and we have therefore been led to consider simpler structures in which the faults appear at regular intervals. In part II, Dr Wilson has shown that the general case of a random distribution of faults can be solved, and this really supersedes our work; nevertheless, we feel that the methods we have used are worth recording since the results were rather unexpected, and may be applicable to other problems.

Consider a structure in which a fault appears after each N perfect unit cells in the c direction. To make this periodic we must have a regular sequence of faults, and the simplest arrangement which contains all possible types in equal proportions is

$$A_1B_2\ldots A_1B_2C_1B_2\ldots C_1B_2C_1A_2\ldots C_1A_2B_1A_2\ldots B_1A_2B_1C_2\ldots B_1C_2A_1C_2\ldots A_1C_2A_1B_2,$$

where the suffixes 1 and 2 represent atoms at heights of 0 and $\frac{1}{2}$ respectively in the basic unit cell. The atomic positions are shown in table 1.

TABLE 1. POSITIONS OF THE ATOMS IN THE LARGE UNIT CELL

type of position	x/a	y/b	z/c†
A_1	0	0	$-N,\ -(N-1),\ \ldots,\ 0,\ \ldots,\ (N-1)$
A_2	0	0	$(2N+\frac{1}{2}),\ (2N+1\frac{1}{2}),\ \ldots,\ (4N-\frac{1}{2})$
B_1	$\frac{1}{3}$	$\frac{2}{3}$	$3N,\ (3N+1),\ \ldots,\ (5N-1)$
B_2	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{1}{2},\ 1\frac{1}{2},\ \ldots,\ (2N-\frac{1}{2})$
C_1	$\frac{2}{3}$	$\frac{1}{3}$	$N,\ (N+1),\ \ldots,\ (3N-1)$
C_2	$\frac{2}{3}$	$\frac{1}{3}$	$(4N+\frac{1}{2}),\ (4N+1\frac{1}{2}),\ \ldots,\ (6N-\frac{1}{2})$

† For convenience c is used instead of C.

The unit cell of this structure is still hexagonal, with the same a- and b-axes as the basic unit cell, but the c-axis, C, is $6N$ times as long. The reciprocal lattice would therefore contain points very close together on lines of constant h, k; this is a first step towards the representation of the structure by the reciprocal lattice shown in figure 3. The exact representation may be found by calculating the structure factors for different orders hkL. The index L is related to the corresponding index l for the basic unit cell by the relation

$$L = 6Nl. \tag{1}$$

Since L may have all integral values, l may be non-integral.

The structure factor is

$$F = f\Sigma \exp\left[2\pi i\left(\frac{hx}{a}+\frac{ky}{b}+\frac{Lz}{C}\right)\right], \tag{2}$$

where f is the scattering factor for one atom. The summation is made over all atoms, of positions x, y, z, in the unit cell. It can be seen from table 1 that each particular type of atom forms a series with $2N$ terms, the z co-ordinates of the B_2 atoms, for instance, being $\frac{(2n+1)\,C}{12N}$, where $0 \leqslant n \leqslant 2N-1$. Thus the contribution to the structure factor of the B_2 atoms is

$$f \sum_{n=0}^{2N-1} \exp\left[2\pi i\left(\tfrac{1}{3}h + \tfrac{2}{3}k + \frac{2n+1}{12N}\,L\right)\right]. \tag{3}$$

This is a geometric series of which the sum is

$$f \exp\left[2\pi i\left(\tfrac{1}{3}h + \tfrac{2}{3}k + \frac{1}{12N}\,L\right)\right] \frac{1-\exp\left[2\pi i\,\tfrac{1}{3}L\right]}{1-\exp\left[2\pi i\,\dfrac{1}{6N}\,L\right]}. \tag{4}$$

The series for the other atoms are all similar to this except for the initial terms; thus the latter part of the expression (4) is a common factor of the whole, and the complete expression for F is

$$\begin{aligned} f\,\frac{1-\exp\left[2\pi i\,\tfrac{1}{3}L\right]}{1-\exp\left[2\pi i\,\dfrac{1}{6N}\,L\right]} & \left\{\exp\left[-2\pi i\,\tfrac{1}{6}L\right] + \exp\left[2\pi i\left(\tfrac{1}{3}h + \tfrac{2}{3}k + \frac{1}{12N}\,L\right)\right]\right. \\ & + \exp\left[2\pi i\left(\tfrac{2}{3}h + \tfrac{1}{3}k + \tfrac{1}{6}L\right)\right] + \exp\left[2\pi i\left(\frac{1}{12N}\,L + \tfrac{2}{6}L\right)\right] \\ & \left. + \exp\left[2\pi i\left(\tfrac{1}{3}h + \tfrac{2}{3}k + \tfrac{3}{6}L\right)\right] + \exp\left[2\pi i\left(\tfrac{2}{3}h + \tfrac{1}{3}k + \frac{1}{12N}\,L + \tfrac{4}{6}L\right)\right]\right\}. \end{aligned} \tag{5}$$

This common factor is the same as the expression which appears in the theory of the diffraction grating. Its modulus is

$$\frac{\sin \pi\,\tfrac{1}{3}L}{\sin \pi\,\dfrac{1}{6N}\,L}, \tag{6}$$

which has large maxima only when $L = 6Nm$, where m is an integer. The minima nearest to these peaks occur when $L = 6Nm \pm 3$, and it is well known that no appreciable intensity occurs outside these values. From equation (1) it is seen that the peaks are just those which would occur for the basic cell, and the effect of the faults we have postulated is to give a possibility of each reflexion's having four satellites. The intensities may be found by evaluating expression (5) for the different combinations of indices. There are two essentially different cases.

Case I. $\tfrac{1}{3}(h-k)$ *integral.*

The term in brackets in expression (5) reduces to

$$\frac{1-\exp\left[2\pi i L\right]}{1-\exp\left[2\pi i\,\tfrac{1}{3}L\right]}\left(\exp\left[-2\pi i\,\tfrac{1}{6}L\right] + \exp\left[2\pi i\,\frac{1}{12N}\,L\right]\right)$$

so that we have

$$F = f\frac{1-\exp[2\pi iL]}{1-\exp\left[2\pi i\frac{1}{6N}L\right]}\left(\exp[-2\pi i\tfrac{1}{6}L]+\exp\left[2\pi i\frac{1}{12N}L\right]\right). \qquad (7)$$

This is zero for all values of L except those which are even multiples of $6N$, for when it is an odd multiple the term in brackets is zero. The value of F for all values of L for which it is not zero is $12Nf$, which is the normal value. Thus intensities corresponding to even values of the index l are normal and the odd ones are zero, exactly as for the basic hexagonal unit cell.

Case II. (i) $\frac{1}{3}(h-k)$ = *an integer* + $\frac{1}{3}$; (ii) $\frac{1}{3}(h-k)$ = *an integer* − $\frac{1}{3}$

In this case F is complex, and the square of its amplitude has to be found by multiplying by the complex conjugate. The mathematics is fairly lengthy, but, for both cases, the product may be reduced to

$$\frac{2\sin^2\frac{\pi}{3}L}{\sin^2\frac{\pi}{6N}L}\left\{3+\cos 2\pi\left(\frac{L}{12N}+\frac{L}{2}\right)\left[2\cos 2\pi\left(\pm\frac{1}{3}-\frac{L}{3}\right)+1\right]\right.$$
$$+4\cos 2\pi\left(\pm\frac{1}{3}+\frac{L}{6}\right)\left[\cos 2\pi\frac{L}{2}+\cos 2\pi\frac{L}{12N}\right]$$
$$\left.+2\cos 2\pi\left(\pm\frac{1}{3}-\frac{L}{3}\right)+2\cos 2\pi\frac{L}{12N}\cos 2\pi\frac{L}{2}\right\}. \qquad (8)$$

Now we have already shown that the expression outside the brackets is appreciable only when $L/6N$ is nearly integral. Thus we may put

$$\cos 2\pi L/6N = 1,$$

and so

$$\cos 2\pi L/12N = \pm 1.$$

If we take the plus sign, we find that the expression in the brackets is zero for all values of L except $L = 6Nm+2$ for case (i) or $L = 6Nm-2$ for case (ii), when it is 18. If we take the negative sign, the expression is zero for all values of L except $L = 6Nm+1$ for case (i), or $6Nm-1$ for case (ii); in both these cases it is again 18. The values of the intensities depend on the value chosen for N, but if N is not very small $\dfrac{2\sin^2 2\pi/3}{\sin^2 2\pi/6N}$ is of the order of $27N^2/2\pi^2 = 1{\cdot}4N^2$, and $\dfrac{2\sin^2\pi/3}{\sin^2\pi/6N}$ is of the order of $54N^2/\pi^2 = 5{\cdot}4N^2$. Thus in these two cases the normal lines of intensities $36N^2$ and $108N^2$, corresponding to even and odd values of l respectively, are replaced by rather weaker but still sharp lines displaced from their true positions. The displacement depends on N, and since there is no reason to suppose that N is constant throughout a crystal it seems probable that a blurred line would result.

Such reflexions would be represented in reciprocal space by lines parallel to the c-axis, as shown in figure 3. The reciprocal 'points' with l odd should all be the same length, as also should those with l even. Now, for a hexagonal lattice

$$\sin^2\theta = \frac{\lambda^2}{3a^2}(h^2+hk+k^2)+\frac{\lambda^2}{4c^2}l^2, \tag{9}$$

and so

$$\delta l = \delta\theta\,\frac{2c^2\sin 2\theta}{\lambda^2 l}. \tag{10}$$

This relationship can be tested experimentally.

EXPERIMENTAL WORK

The cobalt used had the following analysis: silicon 0·003 %, iron 0·001 %, nickel 0·007 %; traces of copper, calcium, and aluminium were found spectroscopically. It was obtained from the National Physical Laboratory. Some powder was taken from the lump, and was annealed at 1100° C for 5 days, and then at about 380° C for 1 week. A Debye-Scherrer photograph was taken in a 9 cm. camera, with nickel $K\alpha$ radiation monochromatized by reflexion from penta-erythritol. Nickel $K\alpha$ is the shortest K wave-length which can conveniently be used for cobalt (Edwards & Lipson 1941), and the monochromatization was necessary in order that some of the faint lines might be detected. The photograph is shown in figure 1. The line widths were measured in the way described by Jones (1938), and, since the specimen contained about equal proportions of the cubic and hexagonal phases, the lines from the former were used as standards.

TABLE 2. VALUES OF $\frac{2c^2\sin 2\theta}{\lambda^2 l}\beta$ FOR THE BROADENED LINES

		β (radians)				
line	sin 2θ	left	right	mean	$\frac{2c^2\sin 2\theta}{\lambda^2 l}\beta$	x
$10\bar{1}1$	0·779	0·0038	0·0054	0·0046	0·043	0·19
$10\bar{1}2$	0·927	0·0152	0·0177	0·0160	0·089	0·25
$10\bar{1}3$	0·999	0·0125	0·0131	0·0128	0·051	0·13
$20\bar{2}1$	0·969	0·0039	0·0038	0·0038	0·044	0·16
$20\bar{2}2$	0·866	0·0194	0·0212	0·0203	0·105	0·19
$10\bar{1}4$	0·784	0·0269	0·0575	0·0422	0·099	0·10
$20\bar{2}3$	0·355	0·0310	0·0282	0·0296	0·042	0·10

The results of the measurements are shown in table 2; the two sides of the film were measured separately, and it will be seen that the agreement between them is fair except for $10\bar{1}4$ which is so faint that it is barely visible. β is the 'integral width' (Jones 1938). In the last column, in order to enable the calculation in Part II to be completed, is given a quantity which represents the distance in reciprocal space over which the film was measured for the line concerned. This treatment, of course, does not take the shapes of the lines into account, and it was noted that these shapes were not all the same. $10\bar{1}2$, in particular, was different from the other lines, and seemed to consist of a broad line with a sharper one superimposed. At this stage of the work no explanation of this fact can be offered.

The total intensities of the broadened reflexions appeared to be equal to those from the perfect hexagonal structure, but a precise examination could not be carried out owing to the nearness of the lines with $\frac{1}{3}(h-k)$ integral to cubic lines. The intensities of the lines with $\frac{1}{3}(h-k)$ integral are normal and thus could be used as standards, but only 0004 was at all measurable, and that was too faint to give accurate results.

DISCUSSION

It will be seen that the results of the measurements support the conclusion that $\beta \sin 2\theta/l$ is constant for lines with l even and for lines with l odd, and, therefore, that the 'points' in reciprocal space for each of these cases are lines of constant length parallel to the c-axis. The lines with l even are about twice as long as those with l odd, and this might be thought to be connected with the theoretical conclusion that the shifts of the even reflexions are twice the shifts of the odd ones; on this basis the average size of the regions of perfection can be calculated. If half the measured breadth is taken as the average shift of the reflexion, we have for the even lines

$$\delta l = 0{\cdot}049 = \tfrac{2}{6}N,$$

and for the odd lines

$$\delta l = 0{\cdot}023 = \tfrac{1}{6}N.$$

The value of N derived from these expressions is about 7, which means that, on the average, there will be one fault in every 14 planes of atoms. The more exact analysis in Part II shows that this is not far from the truth.

Perhaps a more valuable use for the method, however, lies in its direct application to structures in which faults are known to appear at regular intervals. Such cases have been found by Chao & Taylor (1940), and it is probable that others exist. For instance, in graphite (Taylor & Laidler 1940) several observers have noted the appearance of extra lines of which no complete explanation has been given. Since graphite has a structure which may suffer from the same kind of faults as cobalt, it may be possible to explain these lines by methods similar to those described in this paper.

We wish to thank Professor Sir Lawrence Bragg, F.R.S. and Dr A. J. Bradley, F.R.S. for their interest in this work, and for their encouragement. Acknowledgement must also be made of the award of a Bathurst Studentship from Newnham College, and a grant from the Department of Scientific and Industrial Research (O. S. E.), and of a grant from the British Electrical and Allied Industries Research Association (H. L.).

REFERENCES

van Arkel, A. E. 1939 *Reine Metalle*, p. 330. Berlin: Springer.
Bragg, W. L. 1933*a* *The crystalline state*, p. 191. London: Bell.
Bragg, W. L. 1933*b* *The crystalline state*, p. 144. London: Bell.
Chao, S. H. & Taylor, W. H. 1940 *Proc. Roy. Soc.* A, **176**, 76.
Edwards, O. S. & Lipson, H. 1941 *J. Sci. Instrum.* **18**, 131.
Hendricks, S. B., Jefferson, M. E. & Shultz, J. F. 1930 *Z. Kristallogr.* **73**, 376.
Jones, F. W. 1938 *Proc. Roy. Soc.* A, **166**, 16.
Jones, F. W. & Sykes, C. 1938 *Proc. Roy. Soc.* A, **166**, 376.
von Laue, M. 1936 *Ann. Phys., Lpz.*, **26**, 55.
Nishiyama, Z. 1936 *Kinzoku no Kenkyu*, **13** (7), 300.
Sykes, C. & Jones, F. W. 1936 *Proc. Roy. Soc.* A, **157**, 213.
Taylor, A. & Laidler, D. 1940 *Nature, Lond.*, **146**, 130.

Imperfections in the structure of cobalt

II. Mathematical treatment of proposed structure

BY A. J. C. WILSON

Cavendish Laboratory, Cambridge

(*Communicated by A. J. Bradley, F.R.S.—Received* 14 *August* 1941)

The close-packed structures are made by piling up close-packed planes of atoms in three different relative positions. In proper hexagonal close-packing only two of the positions are used. The intensity of reflexion of X-rays is calculated for a structure which usually follows the hexagonal scheme, but has occasional faults introducing the third position. It is shown that this model gives a satisfactory explanation of the measured integral breadths of certain reflexions from hexagonal cobalt. The affected reflexions with l even are broadened more than those with l odd. The total intensity of the reflexions is not greatly changed.

The frequency of faults in the specimen investigated by Edwards & Lipson is calculated, and is found to be, on the average, about one in every ten planes.

INTRODUCTION

It is well known (von Laue 1936; Patterson 1939; Ewald 1940) that the reflexions from a structure which builds up in accordance with a space group have a representation in reciprocal space in which the reciprocal lattice 'points' are similar, i.e. identical except in intensity. The 'shape' of the reciprocal points depends on the shape of the crystal. Structures are being found, however, in which the reciprocal points are not similar, two striking examples being hexagonal cobalt (Hendricks, Jefferson & Shultz 1930; Edwards & Lipson 1941) and $AuCu_3$ containing 'anti-phase' domains (Sykes & Jones 1936; Jones & Sykes 1938). The observable effect is that certain Debye-Scherrer lines are sharp and others broad in a way which cannot be represented by similar points in the reciprocal lattice. The simplest explanation seems to be that the structure possesses a lattice in which the unit cells are all of the same size and shape, but that the distribution of atoms within the unit cell varies throughout the crystal. Neighbouring cells tend to be like each other, but at each layer there is a chance of a fault, so that the structure factors of two cells chosen at random may be quite different.

Warren (1941) has considered a more extreme case, that of 'amorphous' graphite. Here the lattice of unit cells as well as the space group is defective, and successive planes of atoms perpendicular to the hexagonal axis are displaced and rotated relative to each other by random amounts.

GENERAL CALCULATION OF THE REFLECTED INTENSITY

Let $\mathbf{r}_j$ be the position of the jth unit cell relative to an origin 0, $\mathbf{p}$ a unit vector parallel to the incident X-ray beam, $\mathbf{q}$ a unit vector parallel to the reflected beam. Then the path difference between rays reflected at the origin and rays reflected at the jth unit cell is $\mathbf{r}_j.\mathbf{p}-\mathbf{r}_j.\mathbf{q} = -\mathbf{r}_j.\mathbf{s}$, where $\mathbf{s}=\mathbf{q}-\mathbf{p}$. Suppose that the sides of the unit cell are $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$. Then the position of the jth unit cell relative to the origin is

$$\mathbf{r}_j = j_1\mathbf{a}_1+j_2\mathbf{a}_2+j_3\mathbf{a}_3, \tag{1}$$

where j_1, j_2, j_3 are integers. The amplitude of the reflected beam is then

$$G = \sum_j F_j \exp(-2\pi i\mathbf{r}_j.\mathbf{s}/\lambda), \tag{2}$$

where
$$F_j = \sum_g f_{gj} \exp(-2\pi i\mathbf{u}_{gj}.\mathbf{s}/\lambda) \tag{3}$$

is the structure factor of the jth unit cell; f_{gj} is the structure factor and $\mathbf{u}_{gj}$ the position of the gth atom in the jth unit cell. The intensity of the reflected beam is

$$I = GG^* = \sum_j\sum_{j'} F_jF^*_{j'} \exp\{-2\pi i(\mathbf{r}_j-\mathbf{r}_{j'}).\mathbf{s}/\lambda\}. \tag{4}$$

Let $j'_1 = j_1+m_1$, $j'_2 = j_2+m_2$, $j'_3 = j_3+m_3$. Equation (4) may be written

$$I = \sum_j\sum_m F_jF^*_{j+m} \exp(2\pi i\mathbf{r}_m.\mathbf{s}/\lambda) = \sum_m\{\sum_j F_jF^*_{j+m}\}\exp(2\pi i\mathbf{r}_m.\mathbf{s}/\lambda), \tag{5}$$

proper attention being paid to the limits of the summations. Consider first the summation over j. In a structure with a perfect space group F_j and F_{j+m} would be equal and $\sum\limits_j F_jF^*_{j+m}$ would be simply

$$(n_1-|m_1|)(n_2-|m_2|)(n_3-|m_3|)FF^*,$$

n_1, n_2, n_3 being the numbers of unit cells in the crystal in the directions of the axes. In structures such as that proposed for cobalt, however, F_j and F_{j+m} are not necessarily equal, and the problem of calculating the intensity of reflexion becomes that of finding the mean value of $F_j F^*_{j+m}$. Let this be J_m. The intensity of reflexion is then

$$I = \sum_m J_m(n_1 - |m_1|)(n_2 - |m_2|)(n_3 - |m_3|)\exp(2\pi i \mathbf{r}_m . \mathbf{s}/\lambda). \qquad (6)$$

THE INTENSITY OF REFLEXION FROM COBALT

The close-packed lattices are made by piling up close-packed planes of atoms like those shown in figure 4 of Part I (Edwards & Lipson 1941). Adjacent planes are displaced relative to each other by $(\mathbf{a}_1 + 2\mathbf{a}_2)/3$ or $(2\mathbf{a}_1 + \mathbf{a}_2)/3$. Let the arrangement with an atom on the edge of the unit cell be called *A*, the others being *B* and *C* respectively. A proper hexagonal structure is made by repeating one of *ABAB*..., *BCBC*..., *CACA*..., *CBCB*..., *BABA*..., *ACAC*... indefinitely. In cobalt it may be assumed that a structure starting *ABAB*... has a tendency to continue thus, but that at each added plane there is a possibility of a fault, and of the sequence continuing *CBCB*... or *ACAC*.... By further faults it is possible for it to become any of the other sequences. The probability, P_m, of the $(j+m)$th plane being like the jth may be obtained in the following way. Suppose that the chance of there being a fault at the $(j+m)$th plane is α. Then

the probability of the $(j+m)$th being like the $(j+m-1)$th is 0,

the probability of the $(j+m)$th being like the $(j+m-2)$th is $1-\alpha$,

the probability of the $(j+m)$th being like neither is α,

so
$$\left.\begin{aligned} P_m &= P_{m-1} \times 0 + P_{m-2}(1-\alpha) + (1 - P_{m-1} - P_{m-2})\,\alpha, \\ P_m &+ \alpha P_{m-1} - (1-2\alpha)\,P_{m-2} = \alpha. \end{aligned}\right\} \qquad (7)$$

The solution of this equation is (Milne-Thomson 1933)

$$P_m = 1/3 + Q_e\rho_e^m + Q_0\rho_0^m, \qquad (8a)$$

where $\rho_e = \{-\alpha + \sqrt{(4 - 8\alpha + \alpha^2)}\}/2, \quad \rho_0 = -\{\alpha + \sqrt{(4 - 8\alpha + \alpha^2)}\}/2, \qquad (8b)$

and Q_e and Q_0 are arbitrary constants. They must be chosen so that $P_0 = 1$, $P_1 = 0$. The values required are readily found to be

$$\left.\begin{aligned} Q_e &= \{1 - (1-\alpha)/\sqrt{(4 - 8\alpha + \alpha^2)}\}/3, \\ Q_0 &= \{1 + (1-\alpha)/\sqrt{(4 - 8\alpha + \alpha^2)}\}/3. \end{aligned}\right\} \qquad (9)$$

Equation (8) has been derived for m positive. It is clear, however, that the chance of the $(j+m)$th plane being like the jth must be the same as the chance of the jth being like the $(j+m)$th, i.e. $P_{-m} = P_m$.

For the present purpose the unit cell may conveniently be taken half the usual length in the direction of the hexagonal axis, so that each unit cell contains one atom only. The possible values of F_j are

$$\left.\begin{aligned} F_a &= 1, \\ F_b &= \exp\{-2\pi i(\mathbf{a}_1+2\mathbf{a}_2).\mathbf{s}/3\lambda\}, \\ \text{and}\qquad F_c &= \exp\{-2\pi i(2\mathbf{a}_1+\mathbf{a}_2).\mathbf{s}/3\lambda\}, \end{aligned}\right\} \tag{10}$$

so that the mean value of $F_jF^*_{j+m}$ is

$$\begin{aligned} J_m &= \{P_mF_a^* + (1-P_m)(F_b^*+F_c^*)/2\}F_a/3 \\ &\quad + \{P_mF_b^* + (1-P_m)(F_c^*+F_a^*)/2\}F_b/3 \\ &\quad + \{P_mF_c^* + (1-P_m)(F_a^*+F_b^*)/2\}F_c/3 \\ &= P_m + (1-P_m)\{\exp[2\pi i(\mathbf{a}_1+2\mathbf{a}_2).\mathbf{s}/3\lambda] + \exp[-2\pi i(\mathbf{a}_1+2\mathbf{a}_2).\mathbf{s}/3\lambda] \\ &\quad + \exp[2\pi i(2\mathbf{a}_1+\mathbf{a}_2).\mathbf{s}/3\lambda] + \exp[-2\pi i(2\mathbf{a}_1+\mathbf{a}_2).\mathbf{s}/3\lambda] \\ &\quad + \exp[2\pi i(\mathbf{a}_1-\mathbf{a}_2).\mathbf{s}/3\lambda] + \exp[-2\pi i(\mathbf{a}_1-\mathbf{a}_2).\mathbf{s}/3\lambda]\}/6. \end{aligned} \tag{11}$$

Let the expression in brackets be $6K$. Then

$$\begin{aligned} J_m &= P_m + (1-P_m)K \\ &= (1+2K)/3 + (1-K)(Q_e\rho_e^m + Q_0\rho_0^m). \end{aligned} \tag{12}$$

Substitution in equation (6) gives

$$\begin{aligned} I = {} & \tfrac{1}{3}(1+2K)\sum_{m_1}(n_1-|m_1|)\exp(2\pi i m_1\mathbf{a}_1.\mathbf{s}/\lambda)\sum_{m_2}(n_2-|m_2|) \\ & \times\exp(2\pi i m_2\mathbf{a}_2.\mathbf{s}/\lambda)\sum_{m_3}(2n_3-|m_3|)\exp(\pi i m_3\mathbf{a}_3.\mathbf{s}/\lambda) \\ & + (1-K)Q_e\sum_{m_1}(n_1-|m_1|)\exp(2\pi i m_1\mathbf{a}_1.\mathbf{s}/\lambda)\sum_{m_2}(n_2-|m_2|) \\ & \times\exp(2\pi i m_2\mathbf{a}_2.\mathbf{s}/\lambda)\sum_{m_3}(2n_3-|m_3|)\rho_e^{|m_3|}\exp(\pi i m_3\mathbf{a}_3.\mathbf{s}/\lambda) \\ & + (1-K)Q_0\sum_{m_1}(n_1-|m_1|)\exp(2\pi i m_1\mathbf{a}_1.\mathbf{s}/\lambda)\sum_{m_2}(n_2-|m_2|) \\ & \times\exp(2\pi i m_2\mathbf{a}_2.\mathbf{s}/\lambda)\sum_{m_3}(2n_3-|m_3|)\rho_0^{|m_3|}\exp(\pi i m_3\mathbf{a}_3.\mathbf{s}/\lambda). \end{aligned} \tag{13}$$

The summation over m_3 contains $(2n_3-|m_3|)$, and 2 is omitted in the exponential part, because the number of planes of atoms in the direction of the hexagonal axis is twice the number of unit cells, and the interplanar spacing is half the edge of the unit cell. The carrying out of the summations is not difficult; the result is

$$\begin{aligned} I = {} & \frac{\sin^2 \pi n_1 \mathbf{a}_1 . \mathbf{s}/\lambda}{\sin^2 \pi \mathbf{a}_1 . \mathbf{s}/\lambda} \frac{\sin^2 \pi n_2 \mathbf{a}_2 . \mathbf{s}/\lambda}{\sin^2 \pi \mathbf{a}_2 . \mathbf{s}/\lambda} \\ & \times \left[\frac{1+2K}{3} \frac{\sin^2 \pi n_3 \mathbf{a}_3 . \mathbf{s}/\lambda}{\sin^2 \pi \mathbf{a}_3 . \mathbf{s}/2\lambda} + (1-K)\, Q_e \left\{ \frac{2n_3(1-\rho_e^2)}{1-2\rho_e \cos \pi \mathbf{a}_3 . \mathbf{s}/\lambda + \rho_c^2} \right. \right. \\ & + 2 \frac{\left[\begin{aligned} & 2\rho_e^2 - \rho_e(1+\rho_e^2) \cos \pi \mathbf{a}_3 . \mathbf{s}/\lambda + \rho_e^{2n_3+1} \cos \pi (2n_3+1)\, \mathbf{a}_3 . \mathbf{s}/\lambda \\ & \quad - 2\rho_e^{2n_3+2} \cos 2\pi n_3 \mathbf{a}_3 . \mathbf{s}/\lambda + \rho_e^{2n_3+3} \cos \pi(2n_3-1)\, \mathbf{a}_3 . \mathbf{s}/\lambda \end{aligned} \right]}{(1-2\rho_e \cos \pi \mathbf{a}_3 . \mathbf{s}/\lambda + \rho_e^2)^2} \Bigg\} \\ & + (1-K)\, Q_0 \left\{ \frac{2n_3(1-\rho_0^2)}{1-2\rho_0 \cos \pi \mathbf{a}_3 . \mathbf{s}/\lambda + \rho_0^2} \right. \\ & \left. \left. + 2 \frac{\left[\begin{aligned} & 2\rho_0^2 - \rho_0(1+\rho_0^2) \cos \pi \mathbf{a}_3 . \mathbf{s}/\lambda + \rho_0^{2n_3+1} \cos \pi (2n_3+1)\, \mathbf{a}_3 . \mathbf{s}/\lambda \\ & \quad - 2\rho_0^{2n_3+2} \cos 2\pi n_3 \mathbf{a}_3 . \mathbf{s}/\lambda + \rho_0^{2n_3+3} \cos \pi(2n_3-1)\, \mathbf{a}_3 . \mathbf{s}/\lambda \end{aligned} \right]}{(1-2\rho_0 \cos \pi \mathbf{a}_3 . \mathbf{s}/\lambda + \rho_0^2)^2} \right\} \right]. \end{aligned} \tag{14}$$

The variation of this expression with $\mathbf{a}_1 . \mathbf{s}/\lambda$ and $\mathbf{a}_2 . \mathbf{s}/\lambda$ is quite normal. The first two factors have the value $n_1^2 n_2^2$ for $\mathbf{a}_1 . \mathbf{s}/\lambda$ and $\mathbf{a}_2 . \mathbf{s}/\lambda$ equal to integers, say h and k, but fall to inappreciable values if $\mathbf{a}_1 . \mathbf{s}/\lambda$ or $\mathbf{a}_2 . \mathbf{s}/\lambda$ differs from an integer by more than $1/n_1$ or $1/n_2$. K is therefore

$$\cos 2\pi(h-k)/3.$$

When $(h-k)$ is divisible by three, $(1+2K)/3$ and $(1-K)$ have the values 1 and 0 respectively, and the expression in square brackets in (14) is also normal, being $4n_3^2$ when $\mathbf{a}_3 . \mathbf{s}/\lambda$ is an even integer, 0 when it is an odd integer. When $(h-k)$ is not divisible by three the expression in square brackets depends in a complicated way on α. If $\alpha \ll 1/n_3$, it reduces to

$$I = \frac{3}{2} \frac{1}{6} \frac{\sin^2 \pi n_3 \mathbf{a}_3 . \mathbf{s}/\lambda}{\sin^2 \pi \mathbf{a}_3 . \mathbf{s}/2\lambda} + \frac{3}{2} \frac{1}{2} \frac{\sin^2 \pi n_3 \mathbf{a}_3 . \mathbf{s}/\lambda}{\cos^2 \pi \mathbf{a}_3 . \mathbf{s}/2\lambda}. \tag{15}$$

This has the value n_3^2 for $\mathbf{a}_3 . \mathbf{s}/\lambda = l$ where l is an even integer, $3n_3^2$ for l an odd integer; i.e. it gives the ordinary reflexions for a hexagonal close-packed structure. When α is 1 the ρ's become complex, and the expression in square brackets reduces to

$$\frac{3}{2} \frac{1}{3} \frac{\sin^2 \pi n_3(\mathbf{a}_3 . \mathbf{s}/\lambda - 2/3)}{\sin^2 \pi(\mathbf{a}_3 . \mathbf{s}/2\lambda - 1/3)} + \frac{3}{2} \frac{1}{3} \frac{\sin^2 \pi n_3(\mathbf{a}_3 . \mathbf{s}/\lambda + 2/3)}{\sin^2 \pi(\mathbf{a}_3 . \mathbf{s}/2\lambda + 1/3)}, \tag{16}$$

which is $2n_3^2$ for $\mathbf{a}_3 . \mathbf{s}/\lambda$ equal to an odd integer plus or minus one-third. Consideration shows that these are the positions of reflexions from the two possible face-centred cubic structures with [111] axes in the direction of the original hexagonal axis. A hexagonal structure with a fault at every plane would be one of these cubic structures; the meaning of the result obtained is that one is as probable as the other.

The range of α which is of interest for the explanation of the broad lines of hexagonal cobalt is $1/n_3 \ll \alpha < 1/2$. In this region the important terms in the square brackets are

$$\frac{3}{2}\left\{Q_e\frac{}{1-2\rho_e\cos\pi l+\rho_e^2}+Q_0\frac{}{1-2\rho_0\,\mathrm{c}}\right.$$

where l has been written for $\mathbf{a}_3 . \mathbf{s}/\lambda$, though the intensity is appreciable for l differing considerably from an integer. The first term has a maximum of $3n_3 Q_e(1+\rho_e)/(1-\rho_e) \risingdotseq 2n_3/3\alpha$ for l an even integer, the second a maximum of $3n_3 Q_0(1-\rho_0)/(1+\rho_0) \risingdotseq 6n_3/\alpha$ for l an odd integer. To get the integral breadth (Jones 1938) of these reflexions it is necessary to divide the total intensity by the intensity at the maximum. A simple expression is obtained if each reflexion is taken as extending from $l-1$ to $l+1$, where l is an integer at which the intensity is maximal. The total intensities of the reflexions are

$$\begin{aligned} T &= 3Qn_3\int_{l-1}^{l+1}\frac{(1-\rho^2)\,dl}{1-2\rho\cos\pi l+\rho^2} \\ &= 6Qn_3. \end{aligned} \tag{18}$$

For small values of α, Q_e is $1/6$ and Q_0 is $1/2$ (within 2·5 % for $\alpha < 0{\cdot}2$), so the total intensities are approximately n_3 for l even and $3n_3$ for l odd. The corresponding values for a perfect structure are exactly n_3 and $3n_3$, so the total intensities are right. The integral breadths of the broadened lines are then

$$M = 6Qn_3\Big/3Qn_3\frac{1+\rho}{1-\rho} = 2(1-\rho)/(1+\rho), \tag{19}$$

where ρ is ρ_e for l even, $-\rho_0$ for l odd. The even lines are considerably broader than the odd lines, for

$$M_e/M_0 = (1-\rho_e)(1-\rho_0)/(1+\rho_e)(1+\rho_0) = 3. \tag{20}$$

These values of T, M, and M_e/M_0 are valid only when α is so small that practically all the reflected intensity is got by photometry. It is of course impossible to photometer over the range $l-1$ to $l+1$. Ideally it would be possible to photometer from $l-0{\cdot}5$ to $l+0{\cdot}5$, but the ranges actually covered in Part I were about $l-0{\cdot}15$ to $l+0{\cdot}15$. It is therefore necessary to derive more complicated formulae which include explicitly the range $l-x$ to $l+x$ over which the photometering was done. The whole area under the curve in figure 1 is included in deriving equations (18) to (20), but in the actual photometry the area shaded is lost. The apparent total intensity is therefore

$$\begin{aligned} T &= \int_{l-x}^{l+x} I(l)\,dl - 2xI(x) \\ &= 3Qn_3\left\{\frac{4}{\pi}\arctan\left(\frac{1+\rho}{1-\rho}\tan\frac{\pi x}{2}\right)-\frac{2x(1-\rho^2)}{1-2\rho\cos\pi x+\rho^2}\right\}, \end{aligned} \tag{21}$$

and the apparent intensity at the maximum is

$$I_{\max.} - I(x) = 3Qn_3\left\{\frac{1+\rho}{1-\rho} - \frac{1-\rho^2}{1-2\rho\cos\pi x+\rho^2}\right\}, \tag{22}$$

where again ρ is ρ_e for l even, $-\rho_0$ for l odd.

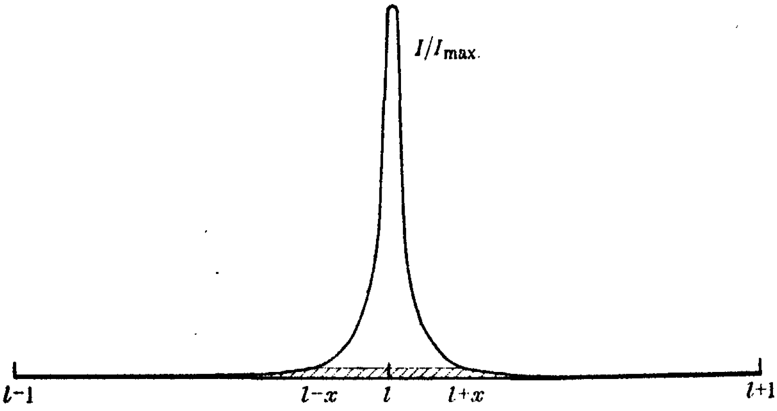


FIGURE 1. The calculated intensity of reflexion as a function of l. The area shaded is lost by the limitations of photometry. The figure is to scale for $\rho = 0{\cdot}9$, $x = 0{\cdot}2$.

The reduction in total intensity predicted by equation (21) is of the order of 15 % for the stronger lines, 30 % for the weaker. Such reductions have not been noted by Edwards & Lipson, but in the absence of reliable comparison lines this is not surprising. They have, however, very kindly allowed me to examine their intensity data. On the present theory (though not on theirs) lines $10\bar{1}0$ and $20\bar{2}0$ as well as 0004 are unaffected in intensity, and may be used as standards. They are relatively stronger than the other lines, and making the calculated corrections improves somewhat the internal consistency of the measurements. Because of the difficulties of measurement of the broad weak lines, however, it seems unwise to draw any definite conclusions.

The integral breadths become

$$M = \frac{\dfrac{4}{\pi}\arctan\left(\dfrac{1+\rho}{1-\rho}\tan\dfrac{\pi x}{2}\right) - \dfrac{2x(1-\rho^2)}{1-2\rho\cos\pi x+\rho^2}}{\dfrac{1+\rho}{1-\rho} - \dfrac{1-\rho^2}{1-2\rho\cos\pi x+\rho^2}}. \tag{23}$$

In figure 2, ρ is shown as a function of M for $x = 0{\cdot}1$, $0{\cdot}2$, $0{\cdot}3$, $0{\cdot}4$, $0{\cdot}5$ and $1{\cdot}0$. In table 1 the actual values of M and x from Part I are given in columns 2 and 3, and the values of ρ derived from them by the use of figure 2 in column 4. Column 5 gives α calculated from

$$\alpha = (1-\rho)(1+\rho)/(2+\rho) \tag{24}$$

which is the reverse of equations (8*b*). In equation (24), unlike (19) and (21)–(23), ρ is ρ_e or ρ_0.

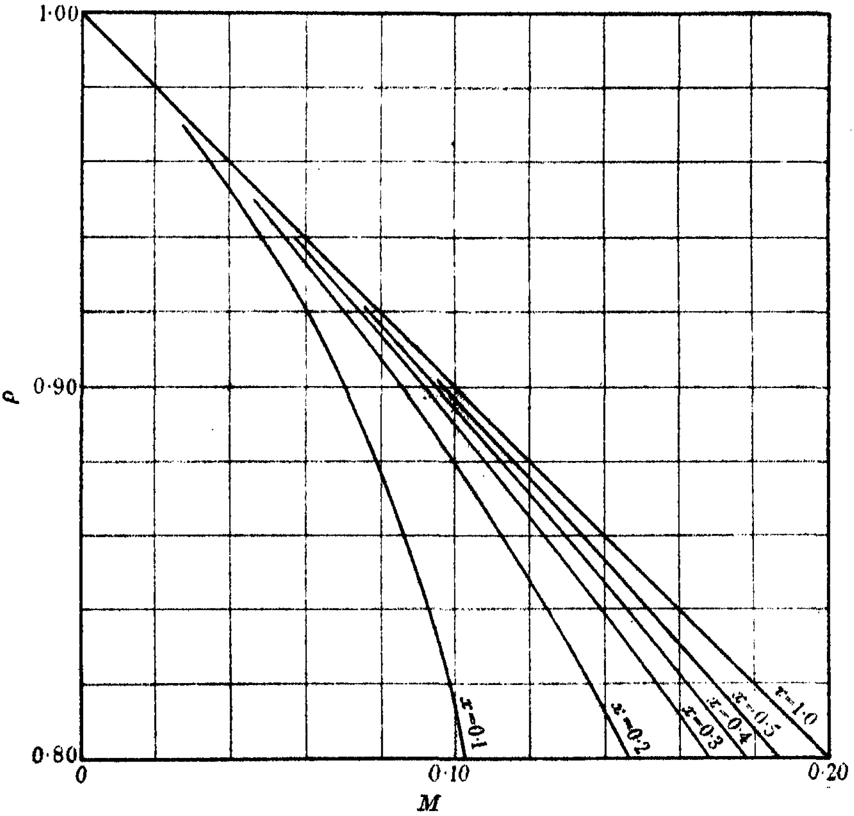


Figure 2. ρ as a function of the integral breadth for $x = 0{\cdot}1, 0{\cdot}2, 0{\cdot}3, 0{\cdot}4, 0{\cdot}5, 1{\cdot}0$.

Table 1. The calculation of α, the probability of a fault

line	M	x	ρ	α	mean deviation
$10\bar{1}1$	0·043	0·19	−0·953	0·088	
$10\bar{1}2$	0·089	0·25	0·899	0·073	
$10\bar{1}3$	0·051	0·13	−0·939	0·111	
$20\bar{2}1$	0·044	0·16	−0·950	0·093	
$20\bar{2}2$	0·105	0·19	0·865	0·088	
$10\bar{1}4$	0·099	0·10	0·815	0·119	
$20\bar{2}3$	0·042	0·10	−0·950	0·093	
mean value				0·095	0·010
mean value from even lines				0·093	0·014
mean value from odd lines				0·096	0·008

The values of α, i.e. the calculated probability of a fault, vary somewhat, but there is no sign of a systematic difference between the lines with l even and the lines with l odd. It seems therefore that the assumption of faults in the cobalt structure, distributed at random along the hexagonal axis, explains satisfactorily both the line-broadening and the difference in breadth between the lines with l even and with l odd.

The author desires to express his thanks to Professor Sir Lawrence Bragg, F.R.S. and Dr A. J. Bradley, F.R.S. for their interest in this work, and to the British Electrical and Allied Industries Research Association for financial assistance.

REFERENCES

Edwards, O. S. & Lipson, H. 1942 *Proc. Roy. Soc.* A, **180**, 268.
Ewald, P. P. 1940 *Proc. Phys. Soc.* **52**, 167.
Hendricks, S. B., Jefferson, M. E. & Shultz, J. F. 1930 *Z. Kristallogr.* **73**, 376.
Jones, F. W. 1938 *Proc. Roy. Soc.* A, **166**, 16.
Jones, F. W. & Sykes, C. 1938 *Proc. Roy. Soc.* A, **166**, 376.
von Laue, M. 1936 *Ann. Phys., Lpz.*, **26**, 55.
Milne-Thomson, L. M. 1933 *The calculus of finite differences*, Chap. XIII. London: Macmillan.
Patterson, A. L. 1939 *Phys. Rev.* **56**, 972.
Sykes, C. & Jones, F. W. 1936 *Proc. Roy. Soc.* A, **157**, 213.
Warren, B. E. 1941 *Phys. Rev.* **59**, 693.

On the theory of finite deformations of elastic crystals

BY R. FÜRTH

Edinburgh

(*Communicated by M. Born, F.R.S.—Received* 10 *September* 1941)

The general theory of finite deformation of cubic crystals at zero temperature is developed to a second-order approximation, and the cases of (1) a uniform hydrostatic pressure, (2) a tension in the direction of one of the axes, (3) a shear along the (0, 1, 0) planes, and (4) a shear along the (0, 1, 1) planes of the lattice, are worked out in detail. A number of 'second-order effects' (deviations from Hooke's law) are predicted which in case (1) have been observed and measured by Bridgman, and in the remaining cases certainly can be detected and measured by suitable experimental arrangements. Assuming the particular force law between the particles of the lattice which was first introduced by Mie and Grüneisen, and later used in the investigations of Lennard-Jones and of Born and his collaborators, and using some of the results of the latter authors, the constants governing the above-mentioned second-order effects are expressed in terms of the constants governing the force law, and calculated numerically for a number of special values of these constants. Thus by comparing the calculated values of these constants with the results of measurements at low temperature the unknown force law could probably be determined.

1. The classical theory of elasticity is based upon two assumptions: that the potential energy of elastic deformation is a quadratic form in the strain components, and that the deformations are infinitely small. An immediate consequence of these assumptions is the validity of Hooke's law, namely, the linear connexion between strain and stress. Actually the deformations are always finite, and therefore also the first assumption, which holds for any reasonable form of potential energy function in the case of infinitely small deformations, is no longer valid and has to be replaced by a more general assumption about this function. Although within the range of purely elastic behaviour of real solids these deformations are in fact generally very small, there are certain cases, for instance, compression of solids under very high uniform pressure, where the deformations are large enough to produce marked deviations from Hooke's law, and also in other cases the accuracy of measurement could undoubtedly be increased so as to detect deviations from Hooke's law within the range of elastic deformations, even if they are comparatively small. The problem of developing a theory of finite elastic deformations of solids is therefore of some physical interest.

The theory of finite deformations of isotropic elastic solids has been developed independently by Murnaghan (1937) and Brillouin (1938). These theories give the general relation between strain and stress for any supposed form of the potential energy function consistent with the assumption of an isotropic medium. If, for example, this function is developed with respect to the strain components up to the third order, instead of up to the second order as in the classical theory, the elastic properties of the medium are determined by five instead of two elastic constants, and the stress components are not linear but quadratic functions of the strain components. By comparison with experimental measurements it should be possible to test whether these relations are satisfied, and to determine the values of the elastic constants. Although this is perhaps interesting from the technical point of view, it is not interesting for the physicist unless the knowledge thus gained can be used to reveal new facts about the molecular structure of solids and the laws governing the forces between their particles. This, however, can only be achieved by developing a theory of finite elastic deformations of crystals. Indeed, for a crystal with given lattice structure and given force law between the particles, the potential energy is a known function of the deformation and the constants appearing in the force law, and therefore the elastic constants of the first and of the second order can be expressed in terms of those constants. Thus by measuring the elastic constants the unknown constants of the force law could be determined.

In the following investigation the theory for cubic crystals at zero temperature is developed, for which a certain amount of work has already been done in a paper by Born & Misra (1940) (further quoted as B.M.). The second-order effects for three particular cases, namely, compression under uniform pressure, tension in the direction of one of the crystallographic axes, and shear along the (0, 1, 0) and the (0, 1, 1) planes of the lattice, have been calculated explicitly for the face-centred cubic lattice (the only stable one as shown by Born (1940) and Misra (1940)), and for a special form of force law. For comparison with experiment only Bridgman's high-pressure experiments can be used at present, but other suitable experiments could certainly be performed without great difficulty at low temperatures, the results of which would provide a good test for the theory and the assumed form of the force law.

2. Consider a cubic lattice with lattice constant a, containing N equal particles. Let $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ be the vectors determining the elementary cell of the deformed lattice. Then the position vector $\mathbf{r}^l$ of the lattice point (l_1, l_2, l_3) is

$$\mathbf{r}^l = l_1\mathbf{a}_1 + l_2\mathbf{a}_2 + l_3\mathbf{a}_3, \tag{1}$$

the length r^l of which is given by

$$(r^l)^2 = \sum_{\alpha\beta} (\mathbf{a}_\alpha . \mathbf{a}_\beta)\, l_\alpha l_\beta \quad (\alpha, \beta = 1, 2, 3). \tag{2}$$

The corresponding quantity $\mathring{r}^l$ in the undeformed lattice is given by

$$(\mathring{r}^l)^2 = a^2(l_1^2 + l_2^2 + l_3^2). \tag{3}$$

From (2) and (3) we have

$$(r^l)^2 - (\mathring{r}^l)^2 = a^2 \sum_{\alpha\beta} \xi_{\alpha\beta} l_\alpha l_\beta, \tag{4}$$

with

$$\xi_{\alpha\beta} = \frac{\mathbf{a}_\alpha . \mathbf{a}_\beta}{a^2} - \delta_{\alpha\beta}. \tag{5}$$

The dimensionless quantities $\xi_{\alpha\beta}$, thus defined, are the components of a symmetric tensor which represents the strain in the deformed crystal. It is convenient to use also another notation for them, namely,

$$\left.\begin{aligned} \xi_1 &= \xi_{11}, & \xi_2 &= \xi_{22}, & \xi_3 &= \xi_{33}, \\ \eta_1 &= 2\xi_{23}, & \eta_2 &= 2\xi_{31}, & \eta_3 &= 2\xi_{12}. \end{aligned}\right\} \tag{6}$$

The potential energy U of the crystal is a function of the $\xi_{\alpha\beta}$, and it is assumed that this function can be developed as a power series with respect to these quantities. Taking account of the symmetry of the lattice considered, Born & Misra have performed this development up to the third order. If the energy of the undeformed lattice is taken as zero, it can be written as follows:

$$\begin{aligned} U = {} & \mathfrak{S}(1)\{\xi_1 + \xi_2 + \xi_3\} + \mathfrak{S}(2)\{\xi_1^2 + \xi_2^2 + \xi_3^2\} \\ & + \mathfrak{S}(1, 1)\{2(\xi_2\xi_3 + \xi_3\xi_1 + \xi_1\xi_2) + (\eta_1^2 + \eta_2^2 + \eta_3^2)\} + \mathfrak{S}(3)\{\xi_1^3 + \xi_2^3 + \xi_3^3\} \\ & + 3\mathfrak{S}(2, 1)\{\xi_1^2(\xi_2 + \xi_3) + \xi_2^2(\xi_3 + \xi_1) + \xi_3^2(\xi_1 + \xi_2) + \xi_1(\eta_2^2 + \eta_3^2) + \xi_2(\eta_3^2 + \eta_1^2) \\ & + \xi_3(\eta_1^2 + \eta_2^2)\} + 3\mathfrak{S}(1, 1, 1)\{2\xi_1\xi_2\xi_3 + 2\eta_1\eta_2\eta_3 + \xi_1\eta_1^2 + \xi_2\eta_2^2 + \xi_3\eta_3^2\}. \end{aligned} \tag{7}$$

Here the $\mathfrak{S}$ are symbols for certain lattice sums involving the potential energy ϕ between two particles of the lattice.

If the undeformed lattice is in equilibrium the linear term in (7) must vanish identically; hence the equilibrium requires the relation

$$\mathfrak{S}(1) = 0. \tag{8}$$

It can easily be seen that $\mathfrak{S}(2)$ and $\mathfrak{S}(1, 1)$ are simply connected with the two first-order elastic constants c_{11} and c_{12} of a cubic crystal (in Voigt's notation), where the Cauchy relation $c_{12} = c_{44}$ is satisfied:

$$c_{11} = \frac{8\mathfrak{S}(2)}{V_0}, \quad c_{12} = \frac{8\mathfrak{S}(1,1)}{V_0}. \tag{9}$$

Here V_0 is the volume of the undeformed crystal.

Instead of the two constants c_{11} and c_{12} it is often useful to introduce Young's modulus E and the Poisson constant μ, defined in the usual way for a tension in the direction of one of the axes. They are connected with c_{11} and c_{12} by the formulae

$$E = \frac{(c_{11}-c_{12})(c_{11}+2c_{12})}{c_{11}+c_{12}}, \quad \mu = \frac{c_{12}}{c_{11}+c_{12}}. \tag{10}$$

Solving them with respect to c_{11} and c_{12} we obtain

$$c_{11} = \frac{1-\mu}{(1+\mu)(1-2\mu)}E, \quad c_{12} = \frac{\mu}{(1+\mu)(1-2\mu)}E, \quad c_{11}+2c_{12} = \frac{E}{1-2\mu}. \tag{11}$$

The three remaining sums $\mathfrak{S}(3)$, $\mathfrak{S}(2, 1)$, $\mathfrak{S}(1, 1, 1)$ determine the three new second-order elastic constants, which are defined by

$$\gamma_1 = \frac{12\mathfrak{S}(3)}{V_0}, \quad \gamma_2 = \frac{12\mathfrak{S}(2,1)}{V_0}, \quad \gamma_3 = \frac{12\mathfrak{S}(1,1,1)}{V_0}. \tag{12}$$

The general relation between the components of the stress $T_{\alpha\beta}$ and the strain for finite deformations have been formulated by Murnaghan in the paper quoted, by equating the virtual work done by the applied forces, acting on the body, to the corresponding increase in the potential energy. They are, rewritten in the notation given here,

$$T_{\alpha\beta} = \frac{2}{V}\left(\frac{\partial U}{\partial \xi_{\alpha\beta}} + \sum_{\gamma} \xi_{\alpha\beta}\frac{\partial U}{\partial \xi_{\beta\gamma}}\right). \tag{13}$$

Here V is the actual volume of the deformed body which can be expressed in terms of the $\xi_{\alpha\beta}$ as follows:

$$\left(\frac{V}{V_0}\right)^2 = \begin{vmatrix} \xi_{11}+1 & \xi_{12} & \xi_{13} \\ \xi_{12} & \xi_{22}+1 & \xi_{23} \\ \xi_{13} & \xi_{23} & \xi_{33}+1 \end{vmatrix}. \tag{14}$$

It is again more convenient for the present purpose to introduce the variables (6) in (13) and (14). Doing this and considering U as a function of these variables, using further Voigt's notation for the stress components

$$\left.\begin{aligned} &X_x = T_{11}, && Y_y = T_{22}, && Z_z = T_{33}, \\ &X_y = Y_x = T_{12}, && Y_z = Z_x = T_{23}, && Z_x = X_z = T_{31}, \end{aligned}\right\} \qquad (15)$$

it follows that

$$\left.\begin{aligned} X_x &= \frac{1}{V}\left\{2\frac{\partial U}{\partial \xi_1}(1+\xi_1) + \eta_2\frac{\partial U}{\partial \eta_2} + \eta_3\frac{\partial U}{\partial \eta_3}\right\}, \\ &\dots\dots\dots\dots\dots\dots\dots\dots, \\ &\dots\dots\dots\dots\dots\dots\dots\dots, \\ X_y &= \frac{1}{V}\left\{\frac{\partial U}{\partial \eta_3}(2+\xi_1+\xi_2) + \frac{\eta_3}{2}\left(\frac{\partial U}{\partial \xi_1} + \frac{\partial U}{\partial \xi_2}\right) + \frac{1}{2}\left(\eta_1\frac{\partial U}{\partial \eta_2} + \eta_2\frac{\partial U}{\partial \eta_1}\right)\right\}, \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots, \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots, \end{aligned}\right\} \qquad (16)$$

with

$$\left(\frac{V}{V_0}\right)^2 = \begin{vmatrix} \xi_1+1 & \eta_3/2 & \eta_2/2 \\ \eta_3/2 & \xi_2+1 & \eta_1/2 \\ \eta_2/2 & \eta_1/2 & \xi_3+1 \end{vmatrix}. \qquad (17)$$

Since in (7) the development is carried out only up to the third-order terms in ξ, η all terms of higher order than the second are omitted on the right-hand side of (16); hence one needs only the zero and first-order terms in the development of (17) which accordingly becomes

$$\frac{1}{V} = \frac{1}{V_0}[1 - \tfrac{1}{2}(\xi_1+\xi_2+\xi_3)]. \qquad (18)$$

The equations (16) and (18) together represent the strain-stress relations which reduce to Hooke's law for infinitely small values of ξ and η. If one wants to solve the problem of determining the deformation of a crystal produced by the application of given external forces, one has to solve these equations with respect to the ξ's and η's. The corresponding change in volume can then be found by substituting these values into (17).

3. To proceed further, a definite form for the potential energy ϕ between two particles of the lattice is assumed, in the form

$$\phi(r) = u\frac{nm}{n-m}\left[-\frac{1}{m}\left(\frac{r_0}{r}\right)^m + \frac{1}{n}\left(\frac{r_0}{r}\right)^n\right] \quad (n > m), \qquad (19)$$

where ϕ depends only on the mutual distance r of the particles, and which has been found to represent the actual force law well in many cases (Mie 1903; Grüneisen 1912; Born 1923, 1939; Lennard-Jones 1937). Here n and m are numerical constants, r_0 the equilibrium distance of a pair of isolated particles, and u their dissociation energy.

On this assumption the quantities $\mathfrak{S}$ are, according to Born & Misra (B.M. equations (3·1)–(3·5))* given by

$$\left.\begin{aligned}
\mathfrak{S}(1) &= -\frac{N}{4}u\frac{nm}{n-m}\left\{-\left(\frac{r_0}{a}\right)^m S^{(1)}_{m+2}+\left(\frac{r_0}{a}\right)^n S^{(1)}_{n+2}\right\},\\
\mathfrak{S}(2) &= \frac{N}{16}u\frac{nm}{n-m}\left\{-(m+2)\left(\frac{r_0}{a}\right)^m S^{(2)}_{m+4}+(n+2)\left(\frac{r_0}{a}\right)^n S^{(2)}_{n+4}\right\},\\
\mathfrak{S}(1,1) &= \frac{N}{16}u\frac{nm}{n-m}\left\{-(m+2)\left(\frac{r_0}{a}\right)^m S^{(11)}_{m+4}+(n+2)\left(\frac{r_0}{a}\right)^n S^{(11)}_{n+4}\right\},\\
\mathfrak{S}(3) &= -\frac{N}{96}u\frac{nm}{n-m}\left\{-(m+2)(m+4)\left(\frac{r_0}{a}\right)^m S^{(3)}_{m+6}+(n+2)(n+4)\left(\frac{r_0}{a}\right)^n S^{(3)}_{n+6}\right\},\\
\mathfrak{S}(2,1) &= -\frac{N}{96}u\frac{nm}{n-m}\left\{-(m+2)(m+4)\left(\frac{r_0}{a}\right)^m S^{(21)}_{m+6}+(n+2)(n+4)\left(\frac{r_0}{a}\right)^n S^{(21)}_{n+6}\right\},\\
\mathfrak{S}(1,1,1) &= -\frac{N}{96}u\frac{nm}{n-m}\left\{-(m+2)(m+4)\left(\frac{r_0}{a}\right)^m S^{(111)}_{m+6}+(n+2)(n+4)\left(\frac{r_0}{a}\right)^n S^{(111)}_{n+6}\right\}.
\end{aligned}\right\} \tag{20}$$

The quantities $S^{(j)}_i$ appearing here are symbols for certain lattice sums which have been numerically calculated and tabulated by Born & Misra (B.M. table III), for the simple, the face-centred, and the body-centred cubic lattices. They satisfy the identities

$$\left.\begin{aligned}
3S^{(1)}_{n+2} &= S^{(0)}_n,\\
3S^{(2)}_{n+4}+6S^{(11)}_{n+4} &= S^{(0)}_n,\\
3S^{(3)}_{n+6}+18S^{(21)}_{n+6}+6S^{(111)}_{n+6} &= S^{(0)}_n.
\end{aligned}\right\} \tag{21}$$

The equilibrium condition (8), together with the first of the equations (20), determines the ratio r_0/a of the equilibrium distance to the lattice constant

$$\left(\frac{r_0}{a}\right)^{n-m} = \frac{S^{(1)}_{m+2}}{S^{(1)}_{n+2}}, \tag{22}$$

which has to be substituted into the remaining equations (20).

For the following considerations only the ratios $\mathfrak{S}(2)/\mathfrak{S}(1)$ for all three kinds of cubic lattices and the ratios $\frac{\mathfrak{S}(3)}{\mathfrak{S}(2)}, \frac{\mathfrak{S}(2,1)}{\mathfrak{S}(2)}, \frac{\mathfrak{S}(1,1,1)}{\mathfrak{S}(2)}$ for the face-centred lattice, which are independent of N and u, depending only on the two numbers n and m are required. I have calculated these quantities for a few typical combinations of values of n and m, and the results of this calculation are contained in table 1.

---

* Equation (3·3) of that paper contains a few mistakes; it should read

$$\mathfrak{S}(k_1k_2k_3)=\frac{N}{2^{k+1}}\frac{(-1)^k}{k!}\sum_n n(n+2)\dots(n+2k-2)\frac{\alpha_n}{a_n}S^{(k_1k_2k_3)}_{n+2k}.$$

TABLE 1

		$\mathfrak{S}(2)/\mathfrak{S}(1,1)$					
m	n	simple	face-centred	body-centred	$\dfrac{\mathfrak{S}(3)}{\mathfrak{S}(2)}$	$\dfrac{\mathfrak{S}(2,1)}{\mathfrak{S}(2)}$	$\dfrac{\mathfrak{S}(1,1,1)}{\mathfrak{S}(2)}$
4	5	− 8·560	1·440	1·056	—	—	—
4	6	—	1·486	1·020	− 1·293	− 0·849	0·0708
4	7	—	—	0·988	—	—	—
4	8	—	1·561	0·950	− 1·473	− 0·913	0·0517
4	12	− 13·13	1·680	—	—	—	—
5	6	—	—	0·986	—	—	—
5	10	—	1·664	0·866	—	—	—
6	12	− 15·89	1·750	0·820	− 1·976	− 1·108	0·0194
7	14	—	1·814	0·770	− 2·230	− 1·209	0·0118
large	large	− large	2·000	1·000	− large	− large	0·0000
					$[\mathfrak{S}(3) = 2\mathfrak{S}(1,1)]$		

The figures of table 1 are first used for determining the value of the Poisson constant μ corresponding to the different types of lattices and the different values of the constants in the force law. According to (9) and (10) we have

$$\mu = \frac{\mathfrak{S}(1,1)}{\mathfrak{S}(2)+\mathfrak{S}(1,1)} = \frac{1}{1+\mathfrak{S}(2)/\mathfrak{S}(1,1)}, \tag{23}$$

by means of which formula the figures of table 2 were calculated.

From the results of this table it can immediately be seen which of the lattices represented there is stable for small perturbations, for the corresponding stability conditions (Born 1940) are

$$c_{11} - c_{12} > 0, \quad c_{12} > 0, \tag{24}$$

which can by means of (10) be transformed into

$$0 < \mu < \tfrac{1}{2}. \tag{25}$$

TABLE 2

		μ		
m	n	simple	face-centred	body-centred
4	5	− 0·132	0·410	0·486
4	6	—	0·402	0·495
4	7	—	—	0·503
4	8	—	0·390	0·512
4	12	− 0·082	0·373	—
5	6	—	—	0·504
5	10	—	0·375	0·536
6	12	− 0·067	0·364	0·549
7	14	—	0·355	0·565
large	large	0·000	0·333	0·500

It can be seen at once from table 2 that only the face-centred lattice satisfies the conditions (25) for all values of n and m larger than 4. The body-centred lattice, however, is only stable for $m < 5$, $n < 7$. The simple lattice, finally, is never stable. These results are, of course, identical with the results of Born & Misra, expressed, however, in a different way.

The figures of tables 1 and 2 are next used for calculating the ratios γ_1/E, γ_2/E, γ_3/E, that is, the values of the elastic constants of the second order measured in the unit E, for from (9), (11) and (12) can easily be deduced the relations

$$\left.\begin{aligned}\frac{\gamma_1}{E} = \frac{3}{2}\frac{1-\mu}{(1+\mu)(1-2\mu)}\frac{\mathfrak{S}(3)}{\mathfrak{S}(2)}, \quad \frac{\gamma_2}{E} = \frac{3}{2}\frac{1-\mu}{(1+\mu)(1-2\mu)}\frac{\mathfrak{S}(2,1)}{\mathfrak{S}(2)},\\ \frac{\gamma_3}{E} = \frac{3}{2}\frac{1-\mu}{(1+\mu)(1-2\mu)}\frac{\mathfrak{S}(1,1,1)}{\mathfrak{S}(2)}.\end{aligned}\right\} \qquad (26)$$

The figures of table 3 for the face-centred lattice and for a few pairs of values of n and m are obtained in this way. They show that the elastic constants of the second order are by no means small compared with those of the first order, but always in the same order of magnitude.

TABLE 3

m	n	γ_1/E	γ_2/E	γ_3/E	μ
4	6	−4·20	−2·76	0·230	0·402
4	8	−4·41	−2·74	0·154	0·390
6	12	−5·07	−2·85	0·050	0·364
7	14	−5·50	−2·97	0·029	0·355
large	large	−large	−large	0·000	0·333

4. The special case where the crystal is subject to a normal stress X_x, Y_y, Z_z with vanishing X_y, Y_z, Z_x is now treated in detail. By symmetry η_1, η_2, η_3 must also vanish under these conditions, and the first three equations (16) become with the help of (18) (if all terms of higher than the second order in the ξ's are omitted)

$$\left.\begin{aligned}X_x &= \frac{1}{V_0}\frac{\partial U}{\partial \xi_1}(2+\xi_1-\xi_2-\xi_3),\\ Y_y &= \frac{1}{V_0}\frac{\partial U}{\partial \xi_2}(2+\xi_2-\xi_3-\xi_1),\\ Z_z &= \frac{1}{V_0}\frac{\partial U}{\partial \xi_3}(2+\xi_3-\xi_1-\xi_2),\end{aligned}\right\} \qquad (27)$$

while the remaining three equations (16) are identically satisfied.

Taking the expression (7) for U, using (9) and (12) and again omitting all terms of higher than the second order it follows from (27)

$$\left.\begin{aligned}c_{11}\xi_1+c_{12}(\xi_2+\xi_3)+b_1\xi_1^2+b_2\xi_1(\xi_2+\xi_3)+b_3(\xi_2^2+\xi_3^2)+b_4\xi_2\xi_3 &= 2X_x,\\ c_{11}\xi_2+c_{12}(\xi_3+\xi_1)+b_1\xi_2^2+b_2\xi_2(\xi_3+\xi_1)+b_3(\xi_3^2+\xi_1^2)+b_4\xi_3\xi_1 &= 2Y_y,\\ c_{11}\xi_3+c_{12}(\xi_1+\xi_2)+b_1\xi_3^2+b_2\xi_3(\xi_1+\xi_2)+b_3(\xi_1^2+\xi_2^2)+b_4\xi_1\xi_2 &= 2Z_z,\end{aligned}\right\} \qquad (28)$$

where

$$b_1 = \gamma_1 + \frac{c_{11}}{2}, \quad b_2 = 2\gamma_2 + \frac{c_{12}-c_{11}}{2}, \quad b_3 = \gamma_2 - \frac{c_{12}}{2}, \quad b_4 = 2\gamma_3 - c_{12}. \tag{29}$$

In order to determine the deformations due to the applied stress the equations (28) must be solved with respect to ξ_1, ξ_2, ξ_3. Approximate solutions can be found in two successive steps by first omitting all the terms of the second order in the ξ's. Since in this case the equations become linear they can be solved immediately. Using (10) one obtains in this way the well-known first-order approximations

$$\left.\begin{aligned} x_1 &= 2x_x = \frac{2}{E}[X_x - \mu(Y_y + Z_z)], \\ x_2 &= 2y_y = \frac{2}{E}[Y_y - \mu(Z_z + X_x)], \\ x_3 &= 2z_z = \frac{2}{E}[Z_z - \mu(X_x + Y_y)], \end{aligned}\right\} \tag{30}$$

where x_x, y_y, z_z are the normal strain components in Voigt's notation.

A second-order approximation is obtained by putting

$$\xi_1 = x_1 + \epsilon_1, \quad \xi_2 = x_2 + \epsilon_2, \quad \xi_3 = x_3 + \epsilon_3, \tag{31}$$

where ϵ_1, ϵ_2, ϵ_3 are small compared with x_1, x_2, x_3. Substituting the expressions (31) into (28) and omitting all terms containing products of the x's and the ϵ's and all terms of the second order in the ϵ's,

$$\left.\begin{aligned} c_{11}\epsilon_1 + c_{12}(\epsilon_2 + \epsilon_3) + b_1 x_1^2 + b_2 x_1(x_2 + x_3) + b_3(x_2^2 + x_3^2) + b_4 x_2 x_3 &= 0, \\ c_{11}\epsilon_2 + c_{12}(\epsilon_3 + \epsilon_1) + b_1 x_2^2 + b_2 x_2(x_3 + x_1) + b_3(x_3^2 + x_1^2) + b_4 x_3 x_1 &= 0, \\ c_{11}\epsilon_3 + c_{12}(\epsilon_1 + \epsilon_2) + b_1 x_3^2 + b_2 x_3(x_1 + x_2) + b_3(x_1^2 + x_2^2) + b_4 x_1 x_2 &= 0. \end{aligned}\right\} \tag{32}$$

Here the expressions (30) are to be substituted for x_1, x_2, x_3. Whence

$$\left.\begin{aligned} c_{11}\epsilon_1 + c_{12}(\epsilon_2 + \epsilon_3) + \alpha_1 X_x^2 + \alpha_2(Y_y^2 + Z_z^2) + \alpha_3 X_x(Y_y + Z_z) + \alpha_4 Y_y Z_z &= 0, \\ c_{11}\epsilon_1 + c_{12}(\epsilon_3 + \epsilon_1) + \alpha_1 Y_y^2 + \alpha_2(Z_z^2 + X_x^2) + \alpha_3 Y_y(Z_z + X_x) + \alpha_4 Z_z X_x &= 0, \\ c_{11}\epsilon_1 + c_{12}(\epsilon_1 + \epsilon_2) + \alpha_1 Z_z^2 + \alpha_2(X_x^2 + Y_y^2) + \alpha_3 Z_z(X_x + Y_y) + \alpha_4 X_x Y_y &= 0, \end{aligned}\right\} \tag{33}$$

with

$$\left.\begin{aligned} \alpha_1 &= \frac{4}{E^2}\{b_1 - 2\mu b_2 + 2\mu^2 b_3 + \mu^2 b_4\}, \\ \alpha_2 &= \frac{4}{E^2}\{\mu^2 b_1 - \mu(1-\mu)\, b_2 + (1+\mu^2)\, b_3 - \mu b_4\}, \\ \alpha_3 &= \frac{4}{E^2}\{-2\mu b_1 + (1 - \mu + 2\mu^2)\, b_2 - 2\mu(1-\mu)\, b_3 - \mu(1-\mu)\, b_4\}, \\ \alpha_4 &= \frac{4}{E^2}\{2\mu^2 b_1 - 2\mu(1-\mu)\, b_2 - 4\mu b_3 + (1+\mu^2)\, b_4\}. \end{aligned}\right\} \tag{34}$$

The equations (33) are linear in ϵ_1, ϵ_2, ϵ_3. Solving them with respect to these quantities it follows by means of (10)

$$\left.\begin{aligned} \epsilon_1 = \frac{1}{E}\{[-\alpha_1+2\mu\alpha_2]X_x^2+[\mu\alpha_1-(1-\mu)\alpha_2](Y_y^2+Z_z^2) \\ +[-(1-\mu)\alpha_3+\mu\alpha_4]X_x(Y_y+Z_z)+[2\mu\alpha_3-\alpha_4]Y_y Z_z\}, \end{aligned}\right\} \quad (35)$$

and similar expressions for ϵ_2 and ϵ_3.

The system of equations (30) and (35) together with (29) and (34) allow, in accordance with (31), the deformations ξ_1, ξ_2, ξ_3 to be expressed in terms of X_X, Y_Y, Z_Z and the elastic constants of the first and the second order.

The corresponding change δV in volume is according to (17) in the present case given by

$$\frac{\delta V}{V_0} = \frac{V-V_0}{V_0} = \sqrt{\{(1+\xi_1)(1+\xi_2)(1+\xi_3)\}}-1. \quad (36)$$

Substituting here from (31) and developing one obtains to the same degree of approximation as before

$$\frac{\delta V}{V_0} = \tfrac{1}{2}(x_1+x_2+x_3)+\tfrac{1}{2}(\epsilon_1+\epsilon_2+\epsilon_3)+\tfrac{1}{2}(x_1x_2+x_2x_3+x_3x_1)-\tfrac{1}{8}(x_1+x_2+x_3)^2, \quad (37)$$

which by means of (30) and (35) can be transformed into

$$\begin{aligned} \frac{\delta V}{V_0} = \frac{1}{E}(1-2\mu)(X_x+Y_y+Z_z)-\frac{1-2\mu}{E} \\ \times\{(\alpha_1+2\alpha_2)(X_x^2+Y_y^2+Z_z^2)+(2\alpha_3+\alpha_4)(X_xY_y+Y_yZ_z+Z_zX_x)\} \\ -\frac{1}{2E^2}\{(1+4\mu)(X_x^2+Y_y^2+Z_z^2)-2(1+2\mu^2)(X_xY_y+Y_yZ_z+Z_zX_x)\}. \quad (38) \end{aligned}$$

5. The case of a crystal under uniform hydrostatic pressure P is realized by putting

$$X_x = Y_y = Z_z = -P. \quad (39)$$

The change in volume produced by this pressure is, according to (34) and (38),

$$\begin{aligned} \frac{\delta V}{V_0} &= -\frac{3(1-2\mu)}{E}P-\frac{3(1-2\mu)}{2E}(\alpha_1+2\alpha_2+2\alpha_3+\alpha_4)P^2+\frac{3(1-2\mu)^2}{2E^2}P^2 \\ &= -\frac{3(1-2\mu)}{E}P+\frac{3}{2}\left(\frac{1-2\mu}{E}\right)^2P^2\left[1-\frac{4(1-2\mu)}{E}(b_1+2b_2+2b_3+b_4)\right]. \end{aligned}$$

Substituting here the expressions (29) and using the third of the relations (11), one obtains

$$\frac{\delta V}{V_0} = -\frac{3(1-2\mu)}{E}P+\frac{3}{2}\left(\frac{1-2\mu}{E}\right)^2P^2\left[3-\frac{4(1-2\mu)}{E}(\gamma_1+6\gamma_2+2\gamma_3)\right], \quad (40)$$

which expresses the relative change in volume in terms of the pressure P and the elastic constants of the first and the second order.

The compressibility κ is defined by

$$\kappa = -\frac{1}{V_0}\frac{\partial V}{\partial P} = -\frac{1}{V_0}\frac{\partial \delta V}{\partial P}. \tag{41}$$

According to (40) therefore

$$\kappa = \kappa_0(1 - f.\kappa_0 P), \tag{42}$$

with

$$\kappa_0 = \frac{3(1-2\mu)}{E} \tag{43}$$

and

$$f = 1 - \frac{4}{3}\frac{1-2\mu}{E}(\gamma_1 + 6\gamma_2 + 2\gamma_3). \tag{44}$$

Here (43) is the well-known expression for the compressibility in the ordinary elasticity theory, and the second term in (42) represents approximately the change of κ due to the deviation from Hooke's law. This deviation vanishes, of course, for small pressures, and it becomes significant if P/E is not very small compared with unity.

In order to discuss formula (42) more conveniently, it may be written in the form

$$-\frac{1}{\kappa_0^2}\frac{\partial \kappa}{\partial P} = f. \tag{45}$$

Assuming now, as in § 3, the particular force law (19) between the particles of the lattice, the γ's obey the relation (26) and μ is given by (23). Using these relations together with the formulae (20) and (22),

$$\left.\begin{aligned} f &= 1 - 2\frac{\mathfrak{S}(3) + 6\mathfrak{S}(2,1) + 2\mathfrak{S}(1,1,1)}{\mathfrak{S}(2)} \cdot \frac{\mathfrak{S}(2)}{\mathfrak{S}(2) + 2\mathfrak{S}(1,1)} \\ &= 1 + \frac{1}{3}\frac{(n+2)(n+4) - (m+2)(m+4)}{(n+2)-(m+2)} = \frac{m+n+9}{3}. \end{aligned}\right\} \tag{46}$$

Hence (45) turns into

$$-\frac{1}{\kappa_0^2}\frac{\partial \kappa}{\partial P} = \frac{m+n+9}{3}. \tag{47}$$

This relation has been derived before by Grüneisen (1912) from his theory of the solid state. It shows that $\partial\kappa/\partial P$ is proportional to κ_0^2 and that the constant of proportionality is the same for all substances obeying the same force law. The values of this constant according to (46) are contained in table 4 which shows that even a very considerable change in the constants of the force law does not affect the order of magnitude of that constant. Practically the same result also follows from Born's theory of the thermodynamics of crystals in the particular case $m = 6$, $n = 12$ supposed by him (Born 1939; Fürth 1941).

In a previous paper (Fürth 1941) I have checked these theoretical conclusions by comparing them with the results of Bridgman's high-pressure experiments, although these experiments have been done on metals, whereas the force law assumed in Born's and in this paper holds only for lattices of atoms held together by Van der Waals forces (Bridgman 1923–35). The values of f derived from these experiments (and tabulated in table 4 of the paper just quoted) are indeed in the same order of magnitude as the figures in table 4, and their average is 7·4, which indicates that the assumption of Lennard-Jones and Born about the most probable value of the constants in the force law represents roughly the real force law, even in the case of metals.

Formula (44) shows that the dependence of the compressibility on the pressure does not vanish even if all the second-order elastic constants are equal to zero. Although there is every reason to believe that the two terms on the right-hand side of (44) will be of the same order of magnitude for every reasonable molecular force law, this fact in itself is not without interest. It is, of course, due to the increase in density of the substance caused by the compression, which can be neglected if the compression is very small. In the paper quoted above Murnaghan comes to the same conclusion for isotropic bodies, and he suggests that the observed dependence of κ on the pressure might be explained in this simple manner. He compares the values of the compression of Na from Bridgman's experiments with his theoretical values and finds very good agreement.

This argument, however, is open to serious objections. There is, in fact, no unique way of expressing the potential energy of deformation as a quadratic form in the deformation quantities, since these quantities can be defined in many different ways which all coincide for infinitely small deformations but differ from each other for finite deformations. Similar ambiguities arise from the definition of 'potential energy of deformation', namely, whether this quantity refers to unit mass or to unit volume of the undeformed body, or finally to unit volume of the deformed body. For each of these cases the statement, that the potential energy should be a quadratic form in the deformation quantities, must obviously lead to different numerical constants for f. Moreover, the very good agreement between the measurements on Na and Murnaghan's theoretical formula is a mere chance, for the values of f deduced from Bridgman's measurements on other substances are spread over a wide range between 1 and 20 (see e.g. Fürth 1941, table 4, p. 46). On the other hand, the method used in this paper gives definite results, once the force law has been chosen.

6. Consider now the case of a crystal under a tension T in the direction of one of the axes. Under this condition

$$X_x = T, \quad Y_y = Z_z = 0. \tag{48}$$

The equations (30) now take the form

$$x_1 = \frac{2}{E}T, \quad x_2 = x_3 = -\frac{2\mu}{E}T, \tag{49}$$

and the equations (35) turn into

$$\left.\begin{aligned} \epsilon_1 &= \frac{1}{E}[-\alpha_1 + 2\mu\alpha_2]\,T^2, \\ \epsilon_2 = \epsilon_3 &= \frac{1}{E}[\mu\alpha_1 - (1-\mu)\,\alpha_2]\,T^2. \end{aligned}\right\} \tag{50}$$

Hence by means of (31),

$$\left.\begin{aligned} \xi_1 &= \frac{1}{E}\{2T + [-\alpha_1 + 2\mu\alpha_2]\,T^2\}, \\ \xi_2 = \xi_3 &= -\frac{\mu}{E}\left\{2T + \left[-\alpha_1 + \frac{1-\mu}{\mu}\alpha_2\right]T^2\right\}. \end{aligned}\right\} \tag{51}$$

The fractional extension $\delta l_1/l_1$ of the crystal in the direction of T is according to the definition of ξ_1 by (5) and (6)

$$\frac{\delta l_1}{l_1} = \sqrt{(1+\xi_1)} - 1 = \tfrac{1}{2}\xi_1 - \tfrac{1}{8}\xi_1^2 + \ldots, \tag{52}$$

and similarly the fractional contraction $-\delta l_2/l_2$ in a perpendicular direction

$$-\frac{\delta l_2}{l_2} = -\sqrt{(1+\xi_2)} + 1 = -\tfrac{1}{2}\xi_2 + \tfrac{1}{8}\xi_2^2 + \ldots. \tag{53}$$

Substituting here for ξ_1 and ξ_2 from (51),

$$\frac{\delta l_1}{l_1} = \frac{T}{E} + \frac{g_1}{2}\left(\frac{T}{E}\right)^2 \tag{54}$$

with

$$g_1 = (-\alpha_1 + 2\mu\alpha_2)\,E - 1. \tag{55}$$

Defining in analogy with the definition of compressibility a coefficient σ of extension, by

$$\sigma = \frac{1}{l_1}\frac{d\delta l_1}{dT}, \tag{56}$$

one obtains from (54)

$$\sigma = \frac{1}{E}\left(1 + g_1\frac{E}{T}\right) = \sigma_0\left(1 + g_1\frac{T}{E}\right), \tag{57}$$

and finally

$$\frac{1}{\sigma_0^2}\frac{d\sigma}{dT} = g_1. \tag{58}$$

Similarly

$$-\frac{\delta l_2}{l_2} = \mu\left[\frac{T}{E} + \frac{g_2}{2}\left(\frac{T}{E}\right)^2\right] \tag{59}$$

with

$$g_2 = \left(-\alpha_1 + \frac{1-\mu}{\mu}\alpha_2\right)E + \mu. \tag{60}$$

The expressions (54) and (59) reduce to the well-known relations of the classical theory for infinitely small values of T/E. But for finite values of this quantity the second term with $(T/E)^2$ becomes significant and causes a deviation from Hooke's law. The actual ratio μ^* of the extension along T and the contraction across T for a finite deformation is different from Poisson's coefficient μ; to the degree of approximation used in this paper it is found that

$$\mu^* = -\frac{\delta l_2}{\delta l_1} = \mu \frac{1+\dfrac{g_2}{2}\dfrac{T}{E}}{1+\dfrac{g_1}{2}\dfrac{T}{E}} = \mu\left(1+h\frac{T}{E}\right) \tag{61}$$

with

$$h = \frac{g_2-g_1}{2} = \frac{1+\mu}{2}\left(1+\frac{1-2\mu}{\mu}\alpha_2 E\right). \tag{62}$$

By means of (34) and (29) the quantities g_1 and h can be easily expressed in terms of the elastic constants of the first and the second order. With the help of (11)

$$g_1 = -(3+4\mu)+\frac{4}{E}\{-\gamma_1(1-2\mu^3)+6\gamma_2\mu(1-\mu+\mu^2)-6\gamma_3\mu^2\}, \tag{63}$$

is obtained, and similarly

$$h = \frac{1+\mu}{2}\left\{1+\frac{4(1-2\mu)}{\mu E}[\gamma_1\mu^2+(1-2\mu+3\mu^2)\gamma_2-2\gamma_3\mu]\right\}. \tag{64}$$

By assuming the particular force law (19) between the particles of the lattice, the γ's are given by (26), and the above formulae take the form

$$g_1 = -(3+4\mu)+6\frac{1-\mu}{(1+\mu)(1-2\mu)} \times\left\{-(1-2\mu^3)\frac{\mathfrak{S}(3)}{\mathfrak{S}(2)}+6\mu(1-\mu+\mu^2)\frac{\mathfrak{S}(2,1)}{\mathfrak{S}(2)}-6\mu^2\frac{\mathfrak{S}(1,1,1)}{\mathfrak{S}(2)}\right\}, \tag{65}$$

$$h = \frac{1+\mu}{2}+3\frac{1-\mu}{\mu}\left\{\mu^2\frac{\mathfrak{S}(3)}{\mathfrak{S}(2)}+(1-2\mu+3\mu^2)\frac{\mathfrak{S}(2,1)}{\mathfrak{S}(2)}-2\mu\frac{\mathfrak{S}(1,1,1)}{\mathfrak{S}(2)}\right\}. \tag{66}$$

These formulae enable one to calculate the numerical values of g_1 and h for different values of the constants m and n of the force law using the figures of tables 2 and 3. Table 4 contains the results of this calculation again for a face-centred lattice, which are also represented graphically in the figure. h is always negative, thus showing that μ^* is always smaller than μ. This second-order effect increases monotonically with increasing m and n and becomes infinitely large in the limiting case $\mu = 1/3$. The quantity g_1 is also always negative, but it decreases first with increasing m and n, passes a minimum and increases afterwards up to a limiting value $-4\cdot33$ for

$\mu = 1/3$. Hence the first of these second-order effects is most significant for large values of m and n and the second for small values of these constants.

Formula (63) shows that one gets a finite and negative value of g_1 by equating all the γ's to zero. This fact has also been observed before by Murnaghan. Its physical meaning, however, is very obscure for reasons similar to those given above, and the argument by which this author comes to the conclusion that the theory of finite deformation predicts a maximum of the deformation-tension curve is not convincing, since equation (54) is only a second approximation to the real relation which can differ widely from (54) for large tensions. A further suggestion of Murnaghan relating these effects with the phenomenon of 'yielding' of a solid under tension must also be rejected, for this effect occurs outside the region of elastic deformation and is due to the quite different process of plastic deformation.

7. As the next special case we consider a crystal under a shear stress S along the $(0, 1, 0)$ or the $(0, 0, 1)$ planes of the lattice. This implies

$$X_x = Y_y = Z_z = X_y = Z_x = 0, \quad Y_z = S. \tag{67}$$

By symmetry it further follows that

$$\eta_2 = \eta_3 = 0, \quad \xi_2 = \xi_3. \tag{68}$$

The first three of the equations (16) contain linear and quadratic terms in the ξ's but only quadratic terms in η_1. This fact, together with the condition (67), implies further that the ξ's are of the order of magnitude of η_1^2, and that all non-linear terms in the ξ's can be omitted. Thus the factor $1/V$ in the equations (16) can according to (18) be identified with $1/V_0$, and the equations reduce to

$$\frac{\partial U}{\partial \xi_1} = 0, \quad 2\frac{\partial U}{\partial \xi_2} + \eta_1 \frac{\partial U}{\partial \eta_1} = 0, \quad \frac{1}{V_0}\left(2\frac{\partial U}{\partial \eta_1} + \eta_1 \frac{\partial U}{\partial \xi_2}\right) = S. \tag{69}$$

Taking the expressions (7) for U one gets from these equations, by using (9) and (12), and neglecting all terms of higher than the first order in ξ_1, ξ_2 and of higher than the second order in η_1

$$\left.\begin{aligned} &c_{11}\xi_1 + 2c_{12}\xi_2 + \gamma_3 \eta_1^2 = 0, \\ &c_{12}\xi_1 + (c_{11}+c_{12})\,\xi_2 + \left(\gamma_2 + \frac{c_{12}}{2}\right)\eta_1^2 = 0, \\ &S = \frac{c_{12}}{2}\eta_1. \end{aligned}\right\} \tag{70}$$

The last of these equations is identical with the corresponding relation in the classical theory, since $\frac{1}{2}\eta_1$ is identical with the strain component $y_z = z_y$ in Voigt's notation. This shows that Hooke's law for a shear is not affected by second-order effects. Such effects, however, are indicated by the first two of the equations (70), consisting in the production of normal deformations by a pure shear, which vanish in the classical approximation. Solving these equations with respect to ξ_1, ξ_2 we obtain

$$\left.\begin{aligned} \xi_1 &= \eta_1^2 \frac{c_{12}^2 + 2c_{12}\gamma_2 - (c_{11}+c_{12})\gamma_3}{(c_{11}-c_{12})(c_{11}+2c_{12})}, \\ \xi_2 = \xi_3 &= \eta_1^2 \frac{-c_{11}c_{12}/2 - c_{11}\gamma_2 + c_{12}\gamma_3}{(c_{11}-c_{12})(c_{11}+2c_{12})}. \end{aligned}\right\} \tag{71}$$

By means of (71), the last of the equations (70), and the formulae (11) we can easily express the fractional changes of lengths of the crystal in the directions of the axes, produced by the shear S, in terms of the elastic constants, namely,

$$\left.\begin{aligned} \frac{\delta l_1}{l_1} &= \frac{\xi_1}{2} = \frac{S^2}{E^2}k_1, \\ \frac{\delta l_2}{l_2} = \frac{\delta l_3}{l_3} &= \frac{\xi_2}{2} = \frac{S^2}{E^2}k_2, \end{aligned}\right\} \tag{72}$$

with

$$\left.\begin{aligned} k_1 &= 2(1+\mu)(1-2\mu)\left\{1 + \frac{(1+\mu)(1-2\mu)}{\mu^2 E}[2\mu\gamma_2 - \gamma_3]\right\}, \\ k_2 &= 2(1+\mu)(1-2\mu)\left\{-\frac{1-\mu}{2\mu} + \frac{(1+\mu)(1-2\mu)}{\mu^2 E}[-(1-\mu)\gamma_2 + \mu\gamma_3]\right\}. \end{aligned}\right\} \tag{73}$$

The corresponding change in volume is found from equation (17) which for the present case reduces to

$$\frac{\delta V}{V_0} = \sqrt{\left\{(1+\xi_1)\left[(1+\xi_2)^2 - \frac{\eta_1^2}{4}\right]\right\}} - 1 = \frac{\xi_1}{2} + \xi_2 - \frac{\eta_1^2}{8}. \tag{74}$$

This can, by means of (72) and (73), be transformed into

$$\frac{\delta V}{V_0} = \frac{S^2}{E^2}\lambda, \tag{75}$$

with
$$\lambda = \frac{(1+\mu)^2(1-2\mu)^2}{\mu^2}\left\{-\frac{2(1-2\mu)}{E}(2\gamma_2+\gamma_3) - \frac{2\mu}{1+\mu} - \tfrac{1}{2}\right\}. \tag{76}$$

Again, assuming the particular force law (19) between the particles of the lattice, the constants k_1, k_2, and λ can, with the help of (26), be transformed into

$$\left.\begin{aligned} k_1 &= (1+\mu)(1-2\mu)\left\{2 + 3\frac{1-\mu}{\mu^2}\left[2\mu\frac{\mathfrak{S}(2,1)}{\mathfrak{S}(2)} - \frac{\mathfrak{S}(1,1,1)}{\mathfrak{S}(2)}\right]\right\}, \\ k_2 &= \frac{(1-\mu^2)(1-2\mu)}{\mu}\left\{-1 + 3\left[-\frac{1-\mu}{\mu}\frac{\mathfrak{S}(2,1)}{\mathfrak{S}(2)} + \frac{\mathfrak{S}(1,1,1)}{\mathfrak{S}(2)}\right]\right\}, \end{aligned}\right\} \tag{77}$$

and

$$\lambda = \frac{(1+\mu)^2(1-2\mu)^2}{\mu^2}\left\{-3\frac{1-\mu}{1+\mu}\left(2\frac{\mathfrak{S}(2,1)}{\mathfrak{S}(2)} + \frac{\mathfrak{S}(1,1,1)}{\mathfrak{S}(2)}\right) - \frac{2\mu}{1+\mu} - \tfrac{1}{2}\right\}. \tag{78}$$

These formulae were again used for calculating the values of the constants k_1, k_2, λ for a face-centred lattice and for different values of m and n. The results of this calculation are also contained in table 4 and are represented by the corresponding graphs of the figure. It is thus seen that k_1 is always negative and k_2 always positive, and the absolute values of both quantities increase monotonically with increasing m and n. This shows that the dimensions of the crystal in the symmetry plane of the shear stress are increased and that they are diminished in the perpendicular direction. The constant λ is always positive, which means that the volume of the crystal is increasing if subject to a shear stress.

TABLE 4

m	n	μ	f	g_1	h	k_1	k_2	λ	$\bar{k}_1$	$\bar{k}_2$	$\bar{\lambda}$
4	6	0·402	6·33	−11·22	−3·08	−1·743	1·230	0·471	21·26	−7·76	1·83
4	8	0·390	7·00	− 9·08	−3·45	−2·25	1·70	0·79	19·43	−6·45	2·67
6	12	0·364	9·00	− 5·38	−4·69	−3·63	3·16	2·12	17·50	−4·30	5·18
7	14	0·355	10·00	− 4·065	−5·30	−4·48	4·04	2·97	17·45	−3·63	6·53
large		0·333	large	− 4·333	−large	−large	large	large	large	−1·78	large

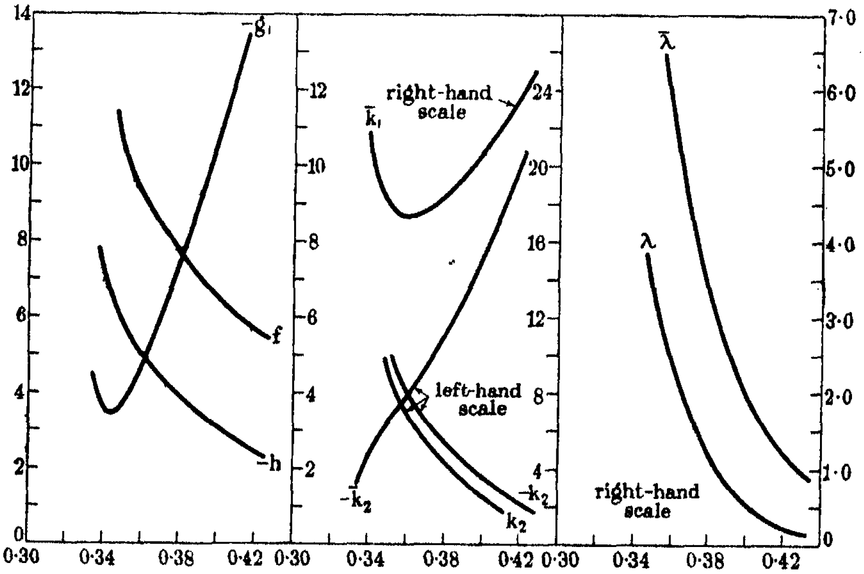


FIGURE 1

8. The general formulae of § 2 of course permit of the determination of the second-order effects for any other direction of a shear stress. Only one example of problems of this kind is given here, namely, the calculation of the second-order effects in a crystal subject to a shear S along the (0, 1, 1) planes of the lattice. Such a shear is equivalent to the superposition of a tension S in the direction of the y-axis and an equal pressure in the direction

of the z-axis (or vice versa). One can therefore use the results of § 4 by putting

$$Y_y = -Z_z = S, \quad X_x = 0. \tag{79}$$

The equations (30) now reduce to

$$\left.\begin{aligned} x_1 &= 0, \\ x_2 &= -x_3 = \frac{2(1+\mu)}{E} S, \end{aligned}\right\} \tag{80}$$

and the equations (35) to

$$\left.\begin{aligned} \epsilon_1 &= \frac{1}{E}[2\mu\alpha_1 - 2(1-\mu)\alpha_2 - 2\mu\alpha_3 + \alpha_4] S^2, \\ \epsilon_2 = \epsilon_3 &= \frac{1}{E}[-(1-\mu)\alpha_1 - (1-3\mu)\alpha_2 + (1-\mu)\alpha_3 - \mu\alpha_4] S^2. \end{aligned}\right\} \tag{81}$$

The relations (34) and (29) can thus be transformed into

$$\left.\begin{aligned} \epsilon_1 &= \frac{8S^2(1+\mu)^2}{E^3}\{\mu(\gamma_1 + c_{11} - c_{12}) - 2(1+\mu)\gamma_2 + 2\gamma_3\}, \\ \epsilon_2 = \epsilon_3 &= \frac{4S^2(1+\mu)^2}{E^3}\{-(1-\mu)(\gamma_1 + c_{11} - c_{12}) + (1+\mu)\gamma_2 - 2\mu\gamma_3\}. \end{aligned}\right\} \tag{82}$$

By means of (80), (82), and (11) the equations (31) now take the form

$$\left.\begin{aligned} \xi_1 &= 2\frac{S^2}{E^2}\bar{k}_1, \\ \xi_2 &= \frac{2(1+\mu)}{E} S + 2\frac{S^2}{E^2}\bar{k}_2, \\ \xi_3 &= -\frac{2(1+\mu)}{E} S + 2\frac{S^2}{E^2}\bar{k}_2, \end{aligned}\right\} \tag{83}$$

where the constants have the following meaning:

$$\left.\begin{aligned} \bar{k}_1 &= 4(1+\mu)\left\{\mu + \frac{1+\mu}{E}[\mu\gamma_1 - (1+\mu)\gamma_2 + \gamma_3]\right\}, \\ \bar{k}_2 &= 2(1+\mu)\left\{-(1-\mu) + \frac{1+\mu}{E}[-(1-\mu)\gamma_1 + (1+\mu)\gamma_2 - 2\mu\gamma_3]\right\}. \end{aligned}\right\} \tag{84}$$

The equations (83) are the analogues of the equations (72), and the constants $\bar{k}_1$, $\bar{k}_2$ play the same part in the present case as the constants k_1, k_2 in the former case.

The change of volume produced by the shear S is found from the formula (37), which by means of (80) and (82) becomes

$$\frac{\delta V}{V_0} = \frac{S^2}{E^2}\bar{\lambda}, \tag{85}$$

with

$$\bar{\lambda} = 2(1+\mu)^2\left\{-\frac{2(1-2\mu)}{E}(\gamma_1 - \gamma_3) - \frac{3(1-\mu)}{1+\mu}\right\}. \tag{86}$$

These formulae are the analogues of the formulae (75) and (76) in the former case.

Finally, introducing the force law (19) as before, the following equations for the constants $\bar{k}_1$, $\bar{k}_2$, $\bar{\lambda}$ are obtained:

$$\left.\begin{aligned} \bar{k}_1 &= 2(1-\mu^2)\left\{\frac{2\mu}{1-\mu}+\frac{3}{1-2\mu}\left[\mu\frac{\mathfrak{S}(3)}{\mathfrak{S}(2)}-(1+\mu)\frac{\mathfrak{S}(2,1)}{\mathfrak{S}(2)}+\frac{\mathfrak{S}(1,1,1)}{\mathfrak{S}(2)}\right]\right\}, \\ \bar{k}_2 &= (1-\mu)^2\left\{-2+\frac{3}{1-2\mu}\left[-(1-\mu)\frac{\mathfrak{S}(3)}{\mathfrak{S}(2)}+(1+\mu)\frac{\mathfrak{S}(2,1)}{\mathfrak{S}(2)}-2\mu\frac{\mathfrak{S}(1,1,1)}{\mathfrak{S}(2)}\right]\right\}, \end{aligned}\right\} \tag{87}$$

$$\bar{\lambda} = 6(1-\mu^2)\left(-\frac{\mathfrak{S}(3)}{\mathfrak{S}(2)}+\frac{\mathfrak{S}(1,1,1)}{\mathfrak{S}(2)}-1\right). \tag{88}$$

These formulae were again used for calculating the values of the constants $\bar{k}_1$, $\bar{k}_2$, and $\bar{\lambda}$ for a face-centred lattice and for a number of pairs of values of the constants m and n in the force law, and the results are tabulated in the last columns of table 4 and represented graphically in the figure. We see that $\bar{\lambda}$ behaves similarly to λ and that both quantities are of the same order of magnitude. Hence it may be concluded that the volume of a cubic crystal will always be increased by the application of a shear stress. The constants $\bar{k}_1$, $\bar{k}_2$, however, behave differently from k_1, k_2, since $\bar{k}_1$ is always positive and $\bar{k}_2$ always negative, contrary to the behaviour of the quantities k_1 and k_2. Moreover, with increasing m and n the value of $\bar{k}_1$ first decreases, passes a minimum and increases again afterwards to infinity as the limit $\mu = 1/3$ is approached; $\bar{k}_2$, on the other hand, decreases monotonically, taking the value $-1{\cdot}78$ in the limit for $\mu = 1/3$. Hence the dimensions of the crystal in the symmetry plane of the shear stress are diminished and the dimensions in the perpendicular direction increased, just opposite to what happens in the case considered in § 7. Thus the second-order effects due to a shear stress depend very markedly on the direction of the stress relative to the axes of the crystal.

If the conditions are slightly modified by reducing the tension in the y-direction by a small amount p, the pressure in the z-direction being kept constant, we have instead of (79)

$$X_x = 0, \quad Y_y = S-p, \quad Z_z = -S. \tag{89}$$

In this case the equations (30) become

$$\left.\begin{aligned} x_1 &= \frac{2}{E}\mu p, \\ x_2 &= \frac{2}{E}[(1+\mu)S-p], \\ x_3 &= -\frac{2}{E}[(1+\mu)S-\mu p]. \end{aligned}\right\} \tag{90}$$

The quantities ϵ_1, ϵ_2, ϵ_3 will remain unchanged as long as p is small compared with S. Formula (37) therefore gives for the change $\delta V'$ in volume by means of (90)

$$\frac{\delta V'}{V_0} = -\frac{1}{E}(1-2\mu)\,p+\frac{\delta V}{V_0}, \tag{91}$$

where $\delta V/V_0$ is given by (85) and (86).

One may now adjust p in such a way that $\delta V'$ vanishes, i.e. that the crystal regains its original volume. This particular value of p is, according to (91), given by

$$\frac{p}{E} = \frac{1}{1-2\mu}\frac{\delta V}{V_0}. \tag{92}$$

Since the right-hand side of this equation is, according to (85), of the order of magnitude S^2/E^2, equation (92) expresses that p/S is of the order of magnitude S/E, that is to say, p is actually small compared with S, in conformity with the above assumption.

The formula (92) provides a possibility for avoiding the difficulty of measuring exactly the increase in volume due to a shear, by measuring instead the diminution of the tension which just restores the original volume of the sample.

REFERENCES

Born, M, 1923 *Atomtheorie des festen Zustandes*, Leipzig.
Born, M. 1939 *J. Chem. Phys.* **7**, 591.
Born, M. 1940 *Proc. Camb. Phil. Soc.* **36**, 160.
Born, M. & Misra, R. D. 1940 *Proc. Camb. Phil. Soc.* **36**, 466.
Bridgman, P. W. 1923–1935 *Proc. Amer. Acad. Arts Sci.*
Brillouin, L. 1938 *Les tenseurs en mécanique et en élasticité*. Paris.
Fürth, R. 1941 *Proc. Camb. Phil. Soc.* **37**, **34**.
Grüneisen, E. 1912 *Ann. Phys., Lpz.*, **39**, **257**.
Lennard-Jones, J. E. 1937 *Physica*, **10**, **941**.
Mie, G. 1903 *Ann. Phys., Lpz.*, **11**, **657**.
Misra, R. D. 1940 *Proc. Camb. Phil. Soc.* **36**, **173**.
Murnaghan, F. D. 1937 *Amer. J. Math.* **59**, **235**.

The effect of thermal vibrations on the scattering of X-rays. II

By Kathleen Sarginson
Edinburgh

(*Communicated by M. Born, F.R.S.—Received* 23 *September* 1941)

The theory of background scattering developed in Part I is worked out in detail. It is shown that only the acoustical branches of the elastic spectrum can produce extra spots. The temperature dependence of the background intensity is discussed. It is shown that the properties of an extra spot in the vicinity of a Laue spot can be derived from the elastic constants of the crystal, and explicit formulae for the position and shape of the spots are given for cubic crystals with diagonal lattices.

Introduction

In part I (Born & Sarginson 1941) the general formulae for the background scattering have been determined, and it has been shown that fairly sharp maxima ought to exist (extra spots). As the general formulae were rather involved, several simplifying assumptions were made, namely, the vectorial elastic waves were replaced by scalar waves and the anisotropy was neglected. Further, the calculation was restricted to the neighbourhood of the Laue spots.

In the present paper some of these restrictions are removed and replaced by better approximations. A method is used, worked out by Waller in his famous dissertation, which is in some way a generalization of Debye's approximation. In the reciprocal space (a Laue spot may be considered as a point of the reciprocal lattice) the waves in the neighbourhood of a Laue spot are not themselves long but correspond to reduced waves of long wavelength, which can be used in a similar way as in the Debye approximation. In this way it can be shown that the background scattering in the neighbourhood of any Laue spot can be expressed in terms of the elastic constants, for any simple Bravais lattice. The intensity function obtained is rather complicated; Waller has not attempted a detailed discussion, but Faxén has derived the existence of secondary maxima (as mentioned in part I). The positions of the maxima will be discussed in two different ways. The first method consists in considering the anisotropy negligible but takes into account the vector character of the waves. Considering the latter influence small, additional terms to our previous formulae are found. The second method considers only Laue spots in the plane normal to the axis of rotation of the crystal. In this case exact formulae are derived for the positions of the extra spots.

It seems to be not worth while to calculate the intensity as long as the simple formulae of part I are not checked by experiment.

1. GENERAL FEATURES OF THE SCATTERING FORMULAE

In Part I it was shown that the intensity of the background scattering was given by (I, (2·41), (2·42), (2·43), (2·44), (2·20))

$$\mathscr{I}_B = \mathscr{I}_0 CN \sum_{kk'} g_k g_{k'} \{\mathscr{B}_{kk'}(Q) \cos 2\pi(\mathbf{s}-\mathbf{s}' . \mathbf{r}_k - \mathbf{r}_{k'}) + i\mathscr{C}_{kk'}(Q) \sin 2\pi(\mathbf{s}-\mathbf{s}' . \mathbf{r}_k - \mathbf{r}_{k'})\},$$

where

$$\mathscr{B}_{kk'}(q) = \sum_j \{\mathscr{L}^j_k(q)\mathscr{L}^{*j}_{k'}(q) + \mathscr{L}^{*j}_k(q)\mathscr{L}^j_{k'}(q)\} \frac{\hbar}{4\omega_j(q)} \coth\frac{\beta_j(q)}{2},$$

$$\mathscr{C}_{kk'}(q) = \sum_j \{\mathscr{L}^j_k(q)\mathscr{L}^{*j}_{k'}(q) - \mathscr{L}^{*j}_k(q)\mathscr{L}^j_{k'}(q)\} \frac{\hbar}{4\omega_j(q)} \coth\frac{\beta_j(q)}{2},$$

$$Q_\alpha = (\mathbf{K}'-\mathbf{K}) . \mathbf{a}_\alpha, \quad \mathscr{L}^j_k(q) = (\mathbf{K}-\mathbf{K}') . \mathbf{e}^j_k(q),$$

$$\beta_j(q) = \hbar\omega_j(q)/kT, \quad g_k = f_k e^{-U_k}, \quad U_k = \tfrac{1}{2}\int_{-\pi}^{+\pi} \mathscr{B}_{kk}(q) \frac{d\bar{q}}{(2\pi)^3}.$$

The symbols used have the following meaning:

$\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$, the lattice vectors defining the smallest possible cell;

$\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$, the reciprocal lattice vectors;

q represents three numbers q_1, q_2, q_3 which are the phases of elastic waves (proportional to the components of the momentum of the elastic quantum in the directions $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$);

$k = 1, 2, \ldots, s$ is the number distinguishing the different atoms in the base;

$j = 1, 2, \ldots, 3s$ is the number distinguishing the different branches of the elastic spectrum;

$\omega_j(q)$, $\mathbf{e}^j_k(q)$ are the proper frequency and the proper vector, respectively, representing the (normalized) amplitude of the atoms of kind k for the normal vibration q in the branch j;

$$\mathbf{K} = \frac{2\pi}{\lambda}\mathbf{s}, \quad \mathbf{K}' = \frac{2\pi}{\lambda}\mathbf{s}'$$

are the wave vectors of the incident and scattered X-rays; λ is the wave-length; $\mathbf{s}$, $\mathbf{s}'$ unit vectors parallel to the incident and reflected wave normal;

f_k is the atomic scattering factor for the atoms of the kind k;

g_k is the scattering factor modified by the thermal agitation.

It is readily seen that $\mathscr{B}_{kk'}(q)$ and $\mathscr{C}_{kk'}(q)$ are quadratic functions in which the coefficients are functions of the eigen-vectors $\mathbf{e}^j_k(q)$ and the eigen-frequencies $\omega_j(q)$ of the elastic lattice vibrations. These quantities are periodic in the lattice. Thus the coefficients in $\mathscr{B}_{kk'}(q)$, $\mathscr{C}_{kk'}(q)$ can be replaced by their reduced values, $\bar{Q}_\alpha = Q_\alpha - 2\pi\kappa_\alpha$ when $2\pi\kappa_\alpha < Q_\alpha < 2\pi(\kappa_\alpha + 1)$,

(κ_1, κ_2, κ_3 integers). This will be convenient later, since it allows a kind of Debye approximation to be made.

I begin with two general considerations concerning the questions which branches of the elastic vibrations are responsible for these extra spots and in which temperature region they are to be expected.

Raman & Nilakantan (1940) have suggested that it is the optical branches which produce the extra spots. It will presently be seen that from the standpoint of our theory the optical branches can produce only a smooth background and no extra spots in the neighbourhood of the Laue spots, and it is the acoustical branch which produces the extra spots.

The extra spots appear in the neighbourhood of Laue spots when the crystal is turned through a small angle from the position giving the Bragg reflexion. As it is these extra spots which interest us we shall consider only the value of $\mathscr{J}_B$ in the neighbourhood of a Laue spot for a setting of the crystal only slightly removed from that which gives rise to a Laue spot. This amounts to treating $\bar{Q}_x$ as small values.

Thus, consideration is given first to the contributions made to $\mathscr{J}_B$ in this region by the optical and acoustical branches of the lattice vibrations. As $q \rightarrow 0$ each of the eigen-vectors $\mathbf{e}^j_k(q)$ approaches some non-zero value. The frequencies of the optical branches of the vibrations, which may be called $\omega_4, \omega_5, \ldots$, can be expanded in the form

$$\omega_j = \omega_j^0 + \omega_j^{(1)} q + \ldots, \quad q^2 = q_1^2 + q_2^2 + q_3^2 \quad (j = 4, 5, \ldots), \tag{1·1}$$

where ω_j^0 are not zero. Thus in the neighbourhood of $q = 0$, $1/\omega_j(q)$ and $\coth(\hbar\omega_j(q)/2kT)$ (for $j = 4, 5, \ldots$) are finite steadily varying functions of q. These have at $q = 0$ (as is well known from the linear model) a maximum of $\omega_j(q)$, hence a minimum of $\dfrac{1}{\omega_j(q)} \coth(h\omega_j(q)/2kT)$. Hence the terms with $j = 4, 5, \ldots$ appearing in $\mathscr{J}_B$ will contribute only a smooth intensity, and cannot give rise to any spots in the neighbourhood of a Laue spot.

However, in the neighbourhood of $q = 0$ the three acoustical branches of the vibrations ω_1, ω_2, ω_3 have the expansions

$$\omega_j(q) = c_j q + \ldots \quad (c_j = \text{velocity of sound}), \tag{1·2}$$

and

$$\coth\frac{\beta_j(q)}{2} = \frac{2kT}{\hbar\omega_j(q)} + \ldots, \quad \frac{1}{\omega_j(q)}\coth\frac{\beta_j(q)}{2} = O(1/q^2). \tag{1·3}$$

Thus the greatest contribution to $\mathscr{J}_B$ arises from this branch of the vibrations. Also any fluctuations in q, although they may be small, will produce comparatively large fluctuations in $1/q^2$, and thus there is a possibility of obtaining fairly well-defined spots. Consequently only the contributions to $\mathscr{J}_B$ arising from the acoustical branches of the vibrations are considered.

It is well known in the theory of lattice vibrations that, as a first approximation, for small q, the eigen-vectors $\mathbf{e}^j_k(q)$, $j = 1, 2, 3$, approach constant real values for all k, i.e. $\mathbf{e}^j(q)$ say. Thus to the first approximation $\mathscr{B}_{kk'}(q)$ and $\mathscr{C}_{kk'}(q)$ approach values independent of k and k' except for a factor $1/\sqrt{(m_k m_{k'})}$, namely,

$$\left.\begin{aligned} &\mathscr{B}_{kk'}(q) \rightarrow \frac{1}{\sqrt{(m_k m_{k'})}} \sum_{j=1}^{3} |\mathscr{L}^j(q)|^2 \frac{\hbar}{2\omega_j(q)} \coth \frac{\beta_j(q)}{2}, \qquad \mathscr{L}^j(q) = (\mathbf{K} - \mathbf{K}') . \mathbf{e}^j(q), \\ &\mathscr{C}_{kk'}(q) \rightarrow 0. \end{aligned}\right\} \tag{1·4}$$

To the same degree everywhere where $\frac{2\pi}{\lambda}(\mathbf{s} - \mathbf{s}')$ appears in $\mathscr{I}_B$ it may be replaced by its value at the corresponding Laue spot, i.e. $2\pi\mathbf{b}_\kappa$:

$$\begin{aligned} \mathscr{I}_B &= \mathscr{I}_0 CN \sum_{kk'} \frac{g_k g_{k'}}{\sqrt{(m_k m_{k'})}} \cos 2\pi (\mathbf{b}_\kappa . \mathbf{r}_k - \mathbf{r}_{k'}) \sum_{j=1}^{3} |\mathscr{L}^j(\bar{Q})|^2 \frac{\hbar}{2\omega_j(Q)} \coth \tfrac{1}{2}\beta_j(\bar{Q}) \\ &\rightarrow \mathscr{I}_0 CN \sum_{kk'} \frac{g_k g_{k'}}{\sqrt{(m_k m_{k'})}} \cos 2\pi (\mathbf{b}_\kappa . \mathbf{r}_k - \mathbf{r}_{k'}) \sum_{j=1}^{3} |\mathscr{L}^j(\bar{Q})|^2 \frac{kT}{c_j^2 \bar{Q}}. \end{aligned} \tag{1·5}$$

For the sake of comparison, the expression for the Laue scattering, $\mathscr{I}_L$, (I (2·40)), is now written down beside this,

$$\mathscr{I}_L = \mathscr{I}_0 CN \sum_{kk'} \frac{g_k g_{k'}}{\sqrt{(m_k m_{k'})}} \cos 2\pi (\mathbf{b}_\kappa . \mathbf{r}_k - \mathbf{r}_{k'}) . (2\pi)^3 \delta(Q), \tag{1·6}$$

were $\delta(q)$ is a periodic δ-function (Laue intensity factor). This shows that the background intensity in the neighbourhood of a Laue spot, and as a consequence that of any extra spot, is proportional to the intensity of the Laue spot.

Consider next how temperature affects the intensities respectively. The temperature in $\mathscr{I}_L$ appears in the factor $g_k g_{k'}$ only. It is in fact $e^{-(U_k + U_{k'})}$, where

$$U_k = \frac{1}{(2\pi)^3} \sum_{j=1}^{3s} \int_{-\pi}^{+\pi} |\mathscr{L}^j_k(q)|^2 \frac{\hbar}{4\omega_j(q)} \coth \left(\frac{\hbar\omega_j(q)}{2kT} \right) d\bar{q}. \tag{1·7}$$

U_k is a positive, increasing function of T. Consequently $e^{-(U_k + U_{k'})}$ decreases as T increases, i.e. the Laue spots are brighter at low temperatures than at high, in agreement with experiment. The part of $\mathscr{I}_B$ which depends on temperature is

$$\sum_j \coth \frac{\hbar\omega_j(q)}{2kT} \sum_{kk'} e^{-(U_k + U_{k'})} / \sqrt{(m_k m_{k'})}. \tag{1·8}$$

Consider next the approximation given in I (2·51) for a monatomic cubic lattice, introducing $\omega_j(q) = \omega_0 w_j(q)$, $\Theta = \frac{\hbar\omega_0}{k}$; ω_0 is the frequency of the oscillator of greatest energy. It follows that the part of the intensity which depends on temperature is

$$\sum_j \coth (\tfrac{1}{2} w_j(\bar{Q}) x) \, e^{-2U(x)}, \quad x = \frac{\Theta}{T}, \tag{1·9}$$

where according to I (2·51)

$$U(x) = \frac{3}{4\pi^2}(1-\cos\vartheta)\frac{\hbar^2}{mk\Theta}\left(\frac{2\pi}{\lambda}\right)^2 \Psi(x). \tag{1·10}$$

We can take for ϑ the angle of deflexion of the neighbouring Laue spot $(\kappa_1, \kappa_2, \kappa_3)$; if $2a$ is the lattice constant we have

$$1-\cos\vartheta = 2\sin^2\vartheta/2 = \kappa^2\left(\frac{\lambda}{a}\right)^2, \quad \kappa^2 = \kappa_1^2+\kappa_2^2+\kappa_3^2,$$

hence $$U(x) = \frac{3}{8}\frac{\hbar^2\kappa^2}{mk\Theta a^2}\Psi(x) = \tfrac{3}{8}\kappa^2\left(\frac{l}{a}\right)^2\Psi(x), \quad l = \frac{\hbar}{\sqrt{(mk\Theta)}}. \tag{1·11}$$

l is a constant of the dimension of length, which has a simple meaning. It is that amplitude of the linear oscillator which corresponds to the maximum energy $k\Theta$. From

$$ml^2\omega_0^2 = \hbar\omega_0, \quad \hbar\omega_0 = k\Theta$$

one gets $$l = \sqrt{\frac{\hbar}{m\omega_0}} = \frac{\hbar}{\sqrt{(mk\Theta)}} = \frac{6{\cdot}97\times10^{-8}}{\sqrt{(\mu\Theta)}} \text{ cm.}, \tag{1·12}$$

where μ is the atomic weight. This quantity l can be called the 'dynamical lattice constant', and U depends on its ratio to the geometrical lattice constant a:

$$U = \Psi(x)/y, \quad y = \frac{8}{3\kappa^2}\left(\frac{a}{l}\right)^2. \tag{1·13}$$

Since only the immediate neighbourhood of a Laue spot is being discussed $\bar{Q}$ is small and thus $w_j(\bar{Q})$ is small. Thus for all temperatures, except those approaching absolute zero, $w_j(\bar{Q})\,x$ is small. Consequently $\coth \tfrac{1}{2}x\,w_j(\bar{Q})$ can be replaced by $2/xw_j(\bar{Q})$, and the function necessary to discuss is (omitting the constant factor $\sum_j 2/w_j(q)$)

$$f(x) = \frac{1}{x}e^{-\Psi(x)/y}. \tag{1·14}$$

Thus $$f'(x) = -\left(\frac{\Psi'(x)}{xy}+\frac{1}{x^2}\right)e^{-\Psi(x)/y}, \tag{1·15}$$

and $f'(x)$ is positive or negative according as

$$y \lesseqgtr -x\Psi'(x). \tag{1·16}$$

This is the criterion determining whether the intensity increases or decreases with decrease of T. Now (I (2·52))

$$\Psi(x) = \frac{1}{4}+\frac{1}{x}\Phi(x), \quad \Phi(x) = \frac{1}{x}\int_0^x \frac{\xi\,d\xi}{e^\xi-1},$$

$$x\Psi'(x) = \frac{1}{e^x-1}-\frac{2}{x}\Phi(x). \tag{1·17}$$

The function $\Phi(x)$ has been discussed in detail by Debye (1914), who has given two expansions valid for large or small x. From these are derived the expansions

$$-x\Psi'(x) = \frac{\pi^2}{3}\frac{1}{x^2} - e^{-x}\left(1+\frac{2}{x}+\frac{2}{x^2}+\dots\right) \quad \text{for large } x, \tag{1·18}$$

and

$$-x\Psi'(x) = \frac{1}{x} - \frac{x}{36} + \dots \quad \text{for small } x. \tag{1·19}$$

Thus for large x, i.e. small temperatures, the expansion on the right-hand side of (1·16) is very small, of the order $1/x^2$. A table of values of $y\kappa^2$ is given at the end of this section (table 1). It is readily seen that for the innermost Laue spots (i.e. κ is not large), which are the ones usually discussed, y is large. Thus the condition $y > -x\Psi'(x)$ is easily fulfilled for low temperatures, and the intensity decreases with decrease of T. For high temperatures the criterion becomes

$$y \lesssim \frac{1}{x}, \quad \text{i.e. } y \lesssim \frac{T}{\Theta}, \tag{1·20}$$

so that the intensity will increase with increase of temperature until the critical value $T_i = \Theta y$ is reached. A table of values of $T_i\kappa^2$ is appended. For all substances T_i is very large, when again consideration is given to Laue spots for which κ is not large.

TABLE 1

substance		μ		Θ °K	$a \times 10^8$ cm.	$y\kappa^2$	$T_i\kappa^2 \times 10^5$ °K
elements:	Ne	20·2		63	2·26	179	0·113
	Ar	39·9		85	2·71	686	0·583
	Li	6·9		510	1·73	290	1·48
	Na	23·0		202	2·12	575	1·162
	K	39·1		126	2·665	964	1·215
	Cu	63·6		315	1·804	1795	5·656
	Ag	107·9		215	2·038	2654	5·707
	Au	197·2		170	2·035	3824	6·501
	Ir	193·1		285	1·915	5560	15·84
	Pt	195·2		226	1·958	4636	10·43
	Pb	207·2		88	2·470	3065	2·70
	C	12·0		2340	1·88	2735	64·0
	Si	28·1		790	2·709	4487	35·45
		μ_1	μ_2				
compounds:	NaCl	23	35·5	281	2·814	2169	6·09
	KCl	39·1	35·5	230	3·138	2326	5·35
	KBr	39·1	79·9	177	3·293	3608	6·39

For very small temperatures (large x) where, in a region surrounding the Laue spot, $\coth \frac{1}{2}xw_j(q)$ cannot be replaced by $2/w_jx$, there is no essential change in the argument. For the function $\Psi(x)$ approaches the constant value $\frac{1}{4}$ for large x, and as y is very large for all substances (see Table 1) $\exp(-\Psi(x)/y)$ is practically equal to unity. The intensity behaves therefore as $\coth(\frac{1}{2}w_j x)$ instead of as $2/w_j x$; it does not fall to zero with decreasing $T = \Theta/x$ but monotonously approaches the unit.

2. GENERAL EXPRESSION FOR THE BACKGROUND SCATTERING

In this section the nature of the background scattering for crystals with the 'diagonal' lattice structure is considered (Born 1923, § 13). Axes fixed in the crystal, Ox, Oy, Oz, which are in the directions of the sides of the cubic cell are chosen, and the unit vectors in these directions are called $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$. Four well-known lattice types which fall into this group are the simple, face-centred and body-centred cubic, and the diamond. Of these, the smallest cell for the first three contains only one particle, and will be defined by the vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$. For the diamond lattice the smallest cell contains two particles. In the theory which has been worked out in part I for the X-ray scattering, it was convenient to refer the crystal to the smallest cell, so that the sums (2·37), (2·38) were taken over all integral values of l_1, l_2, l_3 and thus led immediately to δ-functions. However, in considering the positions of the extra spots it is much more convenient to refer the lattice to the cubic cell.

Call the side of the cubic cell $2a$. Then for the four types of lattice mentioned (considering only these types, as many substances examined are of such structures) the vectors defining the smallest cell are connected with $\mathbf{i}, \mathbf{j}, \mathbf{k}$ by

$$\left.\begin{array}{llll}
\text{Simple:} & & & \\
 & \mathbf{a}_1 = 2a\mathbf{i}, & \mathbf{a}_2 = 2a\mathbf{j}, & \mathbf{a}_3 = 2a\mathbf{k}; \\
\text{Face-centred and diamond:} & & & \\
 & \mathbf{a}_1 = a(\mathbf{j}+\mathbf{k}), & \mathbf{a}_2 = a(\mathbf{k}+\mathbf{i}), & \mathbf{a}_3 = a(\mathbf{i}+\mathbf{j}); \\
\text{Body-centred:} & & & \\
 & \mathbf{a}_1 = a(-\mathbf{i}+\mathbf{j}+\mathbf{k}), & \mathbf{a}_2 = a(\mathbf{i}-\mathbf{j}+\mathbf{k}), & \mathbf{a}_3 = a(\mathbf{i}+\mathbf{j}-\mathbf{k}).
\end{array}\right\} \quad (2\cdot1)$$

The second particle of the basis in the diamond lattice has the position vector $\frac{1}{2}a(\mathbf{i}+\mathbf{j}+\mathbf{k})$.

Denoting by Q'_α the quantity $(\mathbf{K}'-\mathbf{K}).\mathbf{a}_\alpha$ and by Q_1, Q_2, Q_3 the quantities $2a(\mathbf{K}'-\mathbf{K}).\mathbf{i}$, $2a(\mathbf{K}'-\mathbf{K}).\mathbf{j}$, $2a(\mathbf{K}'-\mathbf{K}).\mathbf{k}$ respectively, then the conditions for a Bragg reflexion are

$$Q'_\alpha = 2\pi\kappa_\alpha \quad (\alpha = 1, 2, 3), \qquad (2\cdot2)$$

where κ_α are integers.

The corresponding relations for Q_α are now found. Using the equations given above connecting $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ and $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ the following relations between Q_α and Q'_α are obtained:

$$\left.\begin{array}{l} \text{(i) simple lattice:} \qquad Q_\alpha = Q'_\alpha; \\ \text{(ii) face-centred and diamond:} \\ \quad Q_1 = -Q'_1+Q'_2+Q'_3, \quad Q_2 = Q'_1-Q'_2+Q'_3, \quad Q_3 = Q'_1+Q'_2-Q'_3; \\ \text{(iii) body-centred lattice:} \\ \qquad Q_1 = Q'_2+Q'_3, \quad Q_2 = Q'_3+Q'_1, \quad Q_3 = Q'_1+Q'_2, \end{array}\right\} \qquad (2\cdot3)$$

and the corresponding relations for the Bragg reflexion are

$$\left.\begin{array}{l} \text{(i)} \quad Q_\alpha = 2\pi\kappa_\alpha, \\ \text{(ii)} \quad Q_1 = 2\pi(-\kappa'_1+\kappa'_2+\kappa'_3), \ \ldots, \quad \text{or} \quad Q_\alpha = 2\pi\kappa_\alpha, \\ \text{where } \kappa_\alpha \ (\alpha = 1, 2, 3) \text{ are integers which must be all even or} \\ \quad \text{all odd;} \\ \text{(iii)} \quad Q_1 = 2\pi(\kappa'_2+\kappa'_3), \ \ldots, \quad \text{or} \quad Q_\alpha = 2\pi\kappa_\alpha \\ \text{where } \kappa_\alpha \text{ are integers such that } \kappa_1+\kappa_2+\kappa_3 \text{ is even.} \end{array}\right\} \qquad (2\cdot4)$$

In the vibration equations of the crystal it is most convenient to refer the eigen-vectors to the axes Ox, Oy, Oz. Thus

$$L^j_k(q) = \frac{1}{\sqrt{m_k}}\{(K-K')_x e^j_{kx}(q) + (K-K')_y e^j_{ky}(q) + (K-K')_z e^j_{kz}(q)\},$$

where $(K-K')_x$ is the component of $\mathbf{K}-\mathbf{K}'$ along O_x, and thus

$$(K-K')_x = (\mathbf{K}-\mathbf{K}' \,.\, \mathbf{i}) = -Q_1/2a.$$

Similarly, $\qquad (K-K')_y = -Q_2/2a, \quad (K-K')_z = -Q_3/2a,$

and consequently for a Bragg reflexion

$$(K-K')_x = -\frac{\pi\kappa_1}{a}, \quad (K-K')_y = -\frac{\pi\kappa_2}{a}, \quad (K-K')_z = -\frac{\pi\kappa_3}{a}. \qquad (2\cdot5)$$

The dynamical theory for diagonal lattices has been developed (Born 1923, § 13) when the forces acting between the particles are assumed to be central. The case of the diamond lattice, where, as well as these forces, there act also forces due to the directed valency bonds, has been worked out by Nagendra Nath (1935). I have also developed the theory in a slightly different way, starting from a potential function, and have expressed the equations for the vibrations in terms of the three different elastic constants. The results required in the following were found to hold equally well in the case of central forces as those forces which exist in the case of diamond.

In the last section it was shown that the greatest contribution to $\mathcal{J}_B$ is made by the acoustical branches of the vibrations, and the optical branches can produce only a smooth intensity in the neighbourhood of a Laue spot. Consequently in what follows consideration is given only to the $\omega_j(q)$, $\mathbf{e}^j_k(q)$ corresponding to the acoustical branches of the vibrations, i.e. restricting j to 1, 2, 3. Use is also made of the expansions of $\omega_j(q)$, $\mathbf{e}^j_k(q)$ given by Born (1923, § 16)

$$\left.\begin{aligned} \omega_j^2(q) &= c_j^2 q^2 + \dots, \\ \mathbf{e}^j_k(q) &= \mathbf{e}^j + \mathbf{e}^{j\prime}_k(q) + \dots. \end{aligned}\right\} \tag{2·6}$$

The $\mathbf{e}^j$ are orthogonal and normalized by

$$\sum_j e^j_x e^j_y = \delta_{xy}. \tag{2·7}$$

Then the values of c_j and $\mathbf{e}^j$ are given by the equations (Born 1923, p. 648 (128′))

$$\rho c_j^2 q^2 e^j_x = [(c_{11} - c_{44})\, q_1^2 + c_{44} q^2]\, e^j_x + (c_{12} + c_{44})\, [q_1 q_2 e^j_y + q_1 q_3 e^j_z], \tag{2·8}$$

c_{11}, c_{12}, c_{44} being the only three different elastic constants which can exist for such a type of substance, and ρ the density of the matter.* The eigenvectors are referred to the cubic cell. The three values of c_j^2 are the roots of the secular determinant of these equations.

To the approximation which is being used, the expression for $\mathscr{B}_{kk'}$ is given by (1·4) and that for $\mathscr{I}_B$ by (1·5) which can be written

$$\mathscr{I}_B = \mathscr{I}_0 CN\, A(Q, \bar{Q}) \sum_{kk'} \frac{g_k g_{k'}}{\sqrt{(m_k m_{k'})}} \cos 2\pi\, (\mathbf{b}_\kappa \,.\, \mathbf{r}_k - \mathbf{r}_{k'}), \tag{2·9}$$

where

$$A(Q, \bar{Q}) = \sum_{xy} (K - K')_x\, (K - K')_y\, A_{xy}(\bar{Q}) = \frac{1}{4a^2} \sum_{\alpha\beta} A_{\alpha\beta}(\bar{Q})\, Q_\alpha Q_\beta, \tag{2·10}$$

with
$$A_{xy}(q) = A_{\alpha\beta}(q) = \sum_j e^j_x(q)\, e^j_y(q) / c_j^2 q_j^2 \quad (\alpha, \beta = 1, 2, 3). \tag{2·11}$$

Thus the quantity $A(q)$ must be found. This may be done by using the equations (2·8) and the ortho-normal properties of the $\mathbf{e}^j$. Each of the equations (2·8) is multiplied by $e_{\bar{x}}^{j}/c_j^2 q_j^2$ and summation is made over j. Using the orthogonality conditions (2·7)

---

* When this work was finished a paper of W. H. Zachariasen (*Phys. Rev.* **59**, 860, 1941) appeared, which is a continuation of his previous paper (quoted in part I). It deals with the same problem, but the method is somewhat different from that given here. Zachariasen uses a Born-Kármán model for expressing the eigenvectors with the help of the elastic constants. The derivation given here shows that this is not necessary; the restriction to the neighbourhood of a Laue spot permits the introduction of the elastic constants quite generally without using any model. In view of this and other differences (quite apart from the objections raised to Zachariasen's first paper where Debye's incorrect expansion is used, instead of Ott's method) the publication of the present paper seems well justified.

[*Added in proof*, 28 *January* 1942. Another paper on this subject by H. A. Jahn will appear in these *Proceedings* shortly. At present only a short letter to *Nature* (1941) has been published, but Mr Jahn has kindly informed me about his method. It is in fact the same as that used by Miss Sarginson in the present paper, but numerically developed in another way. Jahn's results are in excellent agreement with the experiments by Lonsdale & Smith. M. BORN.]

$$[(c_{11}-c_{44})\,q_1^2+c_{44}q^2]\,A_{x\bar{x}}+(c_{12}+c_{44})\,(q_1q_2A_{y\bar{x}}+q_1q_3A_{z\bar{x}}) = \rho\delta_{x\bar{x}}, \qquad \bar{x} = x, y, z, \qquad (2{\cdot}12)$$

is obtained immediately, the remaining equations being obtained by cyclic interchange of x, y, z and a corresponding interchange of q_1, q_2, q_3. This set of equations is consistent, and it is easily seen that the $A_{x\bar{x}}$ are ρ times the elements of the matrix which is inverse to

$$\begin{bmatrix} (c_{11}-c_{44})\,q_1^2+c_{44}q^2 & (c_{12}+c_{44})\,q_1q_2 & (c_{12}+c_{44})\,q_1q_3 \\ (c_{12}+c_{44})\,q_1q_2 & (c_{11}-c_{44})\,q_2^2+c_{44}q^2 & (c_{12}+c_{44})\,q_2q_3 \\ (c_{12}+c_{44})\,q_1q_3 & (c_{12}+c_{44})\,q_2q_3 & (c_{11}-c_{44})\,q_3^2+c_{44}q^2 \end{bmatrix}. \qquad (2{\cdot}13)$$

For brevity ϵ, μ are introduced to represent the following combinations of the elastic constants:

$$c_{11}-c_{12}-2c_{44} = \epsilon, \quad c_{12}+c_{44} = \mu; \quad \text{then} \quad c_{11}-c_{44} = \mu+\epsilon. \qquad (2{\cdot}14)$$

Then

$$\frac{A_{xx}}{\epsilon(2\mu+\epsilon)\,q_2^2q_3^2-c_{44}(\mu+\epsilon)\,q_1^2q^2+c_{11}c_{44}q^4} = \frac{A_{yz}}{-\mu q_2q_3(\epsilon q_1^2+c_{44}q^2)}$$

$$= \frac{\rho}{\epsilon^2(3\mu+\epsilon)\,q_1^2q_2^2q_3^2+\epsilon(2\mu+\epsilon)\,c_{44}q^2(q_2^2q_3^2+q_3^2q_1^2+q_1^2q_2^2)+c_{44}^2c_{11}q^6}. \qquad (2{\cdot}15)$$

The other coefficients A_{yy}, ... are obtained by cyclic interchange of the co-ordinates and a corresponding interchange of q_1, q_2, q_3.

3. NEGLIGIBLE ANISOTROPY

For a perfectly isotropic substance ϵ is zero. The positions of the extra spots under the simplifying assumption that the anisotropy is negligible will now be determined. For $\epsilon = 0$ the coefficients are

$$\frac{A_{xx}}{c_{11}q^2-\mu q_1^2} = \frac{A_{yz}}{\mu q_2q_3} = \frac{\rho}{c_{11}c_{44}q^4}. \qquad (3{\cdot}1)$$

Thus the expressions (2·11) become identical, and (2·10)

$$A = \frac{\rho}{4a^2}\,\frac{c_{11}\bar{Q}^2Q^2-\mu(Q_1\bar{Q}_1+Q_2\bar{Q}_2+Q_3\bar{Q}_3)^2}{c_{11}c_{44}\bar{Q}^4}, \qquad (3{\cdot}2)$$

where $Q^2 = Q_1^2+Q_2^2+Q_3^2, \quad \bar{Q}^2 = \bar{Q}_1^2+\bar{Q}_2^2+\bar{Q}_3^2.$

It may be remembered that $Q_\alpha-\bar{Q}_\alpha = 2\pi\kappa_\alpha$, where κ_1, κ_2, κ_3 are integers, representing a point in the reciprocal lattice (Laue spot) and that the $\bar{Q}_\alpha$ can be considered as small numbers. The expression (3·2) has an absolute maximum at $\bar{Q}_\alpha = 0$, i.e. at a Laue spot. For other values of $\bar{Q}_1, \bar{Q}_2, \bar{Q}_3$ it is a positive monotonously decreasing function of $\bar{Q}_1, \bar{Q}_2, \bar{Q}_3$, for $c_{11} > \mu = c_{11} - c_{44}$ and $Q^2 . \bar{Q}^2 > (Q_1\bar{Q}_1 + Q_2\bar{Q}_2 + Q_3\bar{Q}_3)$. If the second term were absent in the numerator the $\bar{Q}$-dependent part would be the same as was found in part I (§§ 3, 4). Since the denominator is of the order of $\bar{Q}^4$ and the numerator only of the order $\bar{Q}^2$, the fluctuations of the expression will be influenced mainly by the fluctuations of the denominator, since $\bar{Q}_1, \bar{Q}_2, \bar{Q}_3$ are small. Thus approximately $\mathscr{I}_B$ will have maxima where $\bar{Q}$, expressed in terms of the small changes in the orientation of the crystal and the direction of the scattered intensity, has minimum values.

Consideration is next given to the critical values of $\bar{Q}$ in the following case: The setting of the crystal is as described in part I. The X-axis is the direction of the incident beam and the crystal is rotated about the Y-axis perpendicular to it, which coincides with the y-axis in the crystal. The setting of the crystal is defined by ϕ, the angle between OX and Ox. The photographic plate is set up in a plane parallel to $x = 0$.

The same notation as given in part I is used. The direction cosines of the scattered intensity are ξ, η, ζ and the values of Q_α are given by I (3·16). A Laue spot is defined by equations I (3·13), which give ξ_κ, η_κ, ζ_κ and ϕ_κ for the indices $\kappa_1, \kappa_2, \kappa_3$. The exact expressions are written down in I (4·1). In the following the suffix κ to the quantities ξ, η, ζ, ϕ is omitted, although it must be remembered that these have the values corresponding to a Bragg reflexion. $\bar{Q}_\alpha$ is obtained from Q_α by giving ξ, η, ζ, ϕ small increments $\bar{\xi}$, $\bar{\eta}$, $\bar{\zeta}$, $\bar{\phi}$, the direction cosines of the scattered intensity being $\xi+\bar{\xi}$, $\eta+\bar{\eta}$, $\zeta+\bar{\zeta}$. The developments of $\bar{Q}_\alpha$ are now

$$\left.\begin{aligned} \bar{Q}_1 &= 2a\frac{2\pi}{\lambda}[\bar{\xi}\cos\phi + \bar{\zeta}\sin\phi + \{(1-\xi)\sin\phi + \zeta\cos\phi\}\bar{\phi} + \ldots], \\ \bar{Q}_2 &= 2a\frac{2\pi}{\lambda}\bar{\eta}, \\ \bar{Q}_3 &= 2a\frac{2\pi}{\lambda}[-\bar{\xi}\sin\phi + \bar{\zeta}\cos\phi + \{(1-\xi)\cos\phi - \zeta\sin\phi\}\bar{\phi} + \ldots], \end{aligned}\right\} \tag{3·3}$$

where $\bar{\xi}$, $\bar{\eta}$, $\bar{\zeta}$ are connected by

$$\xi\bar{\xi} + \eta\bar{\eta} + \zeta\bar{\zeta} = \begin{cases} 0 & \text{to the first order,} \\ -\tfrac{1}{2}(\bar{\xi}^2 + \bar{\eta}^2 + \bar{\zeta}^2) & \text{to the second order.} \end{cases} \tag{3·4}$$

Maintaining only first order terms in (3·3)

$$\bar{Q}^2 = 4a^2\left(\frac{2\pi}{\lambda}\right)^2 \{\bar{\xi}^2 + \bar{\eta}^2 + \bar{\zeta}^2 + \zeta\bar{\xi}\bar{\phi} + (1-\xi)\bar{\zeta}\bar{\phi} + [(1-\xi)^2 + \zeta^2]\bar{\phi}^2 + \ldots], \tag{3·5}$$

is obtained, where one of the variables $\bar{\xi}$, $\bar{\eta}$, $\bar{\zeta}$ has to be eliminated with the help of (3·4), first order terms. The minimum of $\bar{Q}^2$ with this additional condition leads to the same formulae as found in I (4·9), namely,

$$\left.\begin{aligned} \bar{\xi} &= -(1-\xi)\zeta\bar{\phi}, \\ \bar{\eta} &= \eta\zeta\bar{\phi}, \\ \bar{\zeta} &= (\zeta^2+\xi-1)\bar{\phi}; \end{aligned}\right\} \tag{3·6}$$

or expressed in terms of χ, the deflexion of the scattered intensity from the neighbouring Laue spot, and $\bar{\vartheta}$, the change in the angle which the scattered intensity makes with the incident beam,

$$\chi = (1-\xi)\bar{\phi}, \quad \bar{\vartheta} = 2\zeta\tan\frac{\vartheta}{2}\bar{\phi}. \tag{3·6a}$$

Consider now the position of the extra spot when account is taken of the term $\mu(Q_1\bar{Q}_1+Q_2\bar{Q}_2+Q_3\bar{Q}_3)$ in (3·2). The exact value of (3·2), if $\bar{\eta}$ is eliminated with the help of (3·4), is

$$\begin{aligned} A = {} & \frac{\rho}{4a^2c_{11}c_{44}} \\ & \times \frac{c_{11}\left\{\dfrac{1-\zeta^2}{\eta^2}\bar{\xi}^2+\dfrac{1-\xi^2}{\eta^2}\bar{\zeta}^2+\dfrac{2\xi\zeta}{\eta^2}\bar{\xi}\bar{\zeta}+2\zeta\bar{\xi}\bar{\phi}+2(1-\xi)\bar{\zeta}\bar{\phi}+(2-2\xi-\eta^2)\bar{\phi}^2\right\}-\mu\bar{\xi}^2}{\dfrac{1-\zeta^2}{\eta^2}\bar{\xi}^2+\dfrac{1-\xi^2}{\eta^2}\bar{\zeta}^2+\dfrac{2\xi\zeta}{\eta^2}\bar{\xi}\bar{\zeta}+2\zeta\bar{\xi}\bar{\phi}+2(1-\xi)\bar{\zeta}\bar{\phi}+(2-2\xi-\eta^2)\bar{\phi}^2}, \end{aligned} \tag{3·7}$$

and the conditions for a maximum can be written in the following way:

$$\left.\begin{aligned} \frac{1-\xi^2}{\eta^2}\bar{\zeta}+\frac{\zeta\xi}{\eta^2}\bar{\xi}+(1-\xi)\bar{\phi} &= 0, \\ \left(A-\frac{\mu}{c_{11}}\bar{\xi}\right)B-2A\left(B-\frac{\mu}{c_{11}}\bar{\xi}^2\right) &= 0, \end{aligned}\right\} \tag{3·8}$$

where
$$A = \frac{1-\zeta^2}{\eta^2}\bar{\xi}^2+\frac{\zeta\xi}{\eta^2}\bar{\zeta}+\zeta\bar{\phi},$$

$$\begin{aligned} B = {} & \left(\frac{1-\zeta^2}{\eta^2}\bar{\xi}+\frac{\zeta\xi}{\eta^2}\bar{\zeta}+\zeta\bar{\phi}\right)\bar{\xi} \\ & +\left(\frac{1-\xi^2}{\eta^2}\bar{\zeta}+\frac{\xi\zeta}{\eta^2}\bar{\xi}+(1-\xi)\bar{\phi}\right)\bar{\zeta}+[\zeta\bar{\xi}+(1-\xi)\bar{\zeta}+(2-2\xi-\eta^2)\bar{\phi}]\bar{\phi}. \end{aligned}$$

Denote by $\bar{\xi}_0$, $\bar{\zeta}_0$ the solutions (3·6) and let $\bar{\xi}_0+\bar{\xi}_1$, $\bar{\zeta}_0+\bar{\zeta}_1$ be solutions of (3·8). Working on the assumption that $\bar{\xi}_1$, $\bar{\eta}_1$ are small compared with $\bar{\xi}_0$, $\bar{\eta}_0$ only terms of the first order in $\bar{\xi}_1$, $\bar{\eta}_1$ are to be retained in (3·8). If the resulting solution is small it will be clear that our assumption is true and the approximation is valid.

Accordingly, using the expression (3·6) for $\bar{\xi}_0$, $\bar{\zeta}_0$, (3·8) becomes

$$\left.\begin{aligned} &\frac{1-\xi^2}{\eta^2}\bar{\zeta}_1+\frac{\xi\zeta}{\eta^2}\bar{\xi}_1=0,\\ &\frac{\mu}{c_{11}}\left[\bar{\xi}_0\left\{\left(\frac{1-\zeta^2}{\eta^2}\bar{\xi}_1+\frac{\xi\zeta}{\eta^2}\bar{\eta}_1\right)\bar{\xi}_0-(\zeta\bar{\xi}_1+(1-\xi)\bar{\zeta}_1)\bar{\phi}\right\}-\zeta^2(\bar{\xi}_0+\bar{\xi}_1)\bar{\phi}\right]\\ &\qquad\qquad-\left(\frac{1-\zeta^2}{\eta^2}\bar{\xi}_1+\frac{\xi\zeta}{\eta^2}\bar{\zeta}_1\right)\zeta^2\bar{\phi}^2=0.\end{aligned}\right\} \tag{3·9}$$

Eliminating $\bar{\zeta}_1$, $\bar{\xi}_1$ is given by

$$\bar{\xi}_1\left[\frac{\mu}{c_{11}}\left(\frac{1-\zeta\xi}{1+\xi}\right)-\frac{1}{1-\xi^2}\right]=\frac{\mu}{c_{11}}\zeta(1-\xi)\bar{\phi}. \tag{3·10}$$

It is immediately seen that for small ϑ (ϑ is the angle of scattering, $\xi=\cos\vartheta$), the most important term in the coefficient of $\bar{\xi}_1$ is $1/1-\xi^2$, and thus $\bar{\xi}_1$ is of the order of $\zeta(1-\xi)^2$. This is ϑ^5 and is small compared with $\bar{\xi}_0$ which is, by (3·6), of the order of magnitude of ϑ^3. Thus for small angles of scattering, the additional term $\frac{\mu}{c_{11}}\bar{\xi}^2$ in (3·2) makes only a small change in the position of the extra spot. This change, since $\bar{\xi}_1$ is negative, brings the extra spot nearer to the Laue spot.

Determine next the values of χ and the shape of the extra spot when the term $\mu\bar{\xi}^2/c_{11}$ is taken into account in (3·2) by assuming that the approximation previously made is valid. The values of $\bar{\zeta}_1$, $\bar{\eta}_1$ corresponding to the solution found for $\bar{\xi}_1$ are

$$\bar{\zeta}_1=\frac{\mu}{c_{11}}\frac{\zeta^2\xi(1-\xi)}{1-\frac{\mu}{c_{11}}(1-\xi)(1-3\xi)}\bar{\phi},\quad \bar{\eta}_1=\frac{\mu}{c_{11}}\frac{\xi\eta\zeta(1-\xi)}{1-\frac{\mu}{c_{11}}(1-\xi)(1-3\xi)}. \tag{3·11}$$

Thus the position of the extra spot is given by

$$\left.\begin{aligned} \bar{\xi}&=\left\{(\xi-1)\zeta-\frac{\mu}{c_{11}}\frac{\zeta(1-\xi)(1-\xi^2)}{1-\frac{\mu}{c_{11}}(1-\xi)(1-3\xi)}\right\}\bar{\phi}+\dots,\\ \bar{\eta}&=\left\{\eta\zeta+\frac{\mu}{c_{11}}\frac{\xi\eta\zeta(1-\xi)}{1-\frac{\mu}{c_{11}}(1-\xi)(1-3\xi)}\right\}\bar{\phi}+\dots,\\ \bar{\zeta}&=\left\{(\zeta^2+\xi-1)+\frac{\mu}{c_{11}}\frac{\zeta^2\xi(1-\xi)}{1-\frac{\mu}{c_{11}}(1-\xi)(1-3\xi)}\right\}\bar{\phi}+\dots,\end{aligned}\right\} \tag{3·12}$$

and the value of χ^2 is, assuming that ϑ is small, and terms of ϑ^8 and higher orders can be neglected,

$$\chi^2=\left\{(1-\xi)^2+\frac{2\mu}{c_{11}}\frac{\zeta^2\xi^2(1-\xi)^2}{1-\frac{\mu}{c_{11}}(1-\xi)(1-3\xi)}\right\}\bar{\phi}+\dots. \tag{3·13}$$

Thus, now, the position of the extra spot depends not only on the value of ξ for the corresponding Laue spot, but also on the azimuthal angle, which will appear in ζ, i.e. for Laue spots lying on a circle about the incident beam the distance of the extra spot varies with the rotation of the circle. However, the term in which ζ appears is small, of the sixth order in ϑ, and thus the effect is small.

The co-ordinates on the photographic plate, denoted by u, v, are, for small ϑ, to the first order, proportional to η and ζ. Thus the curves of constant intensity are, in the neighbourhood of a Laue spot and taking the Laue spot as origin,

$$\frac{1}{B^2}\left\{B-\frac{\mu}{c_{11}}\left(\frac{\eta u+\zeta v}{\xi}\right)^2\right\} = C, \tag{3·14}$$

where C is a constant which must be less than the maximum value of the left-hand side, and

$$B = \frac{1-\zeta^2}{\xi^2}u^2+\frac{1-\eta^2}{\xi^2}v^2+\frac{2\eta\zeta}{\xi^2}uv-\frac{2\eta\zeta}{\xi}u\bar{\phi}+\frac{2(\eta^2+\xi-1)}{\xi}v\bar{\phi}+(2-2\xi-\eta^2)\bar{\phi}^2. \tag{3·15}$$

When a change of co-ordinates is made corresponding to: (i) a transformation of the origin to the centre of the extra spot as given by (3·6), (ii) a rotation of the axes so that the new axes are along the perpendicular to the meridian through the Laue spot, the equations of the curves of constant intensity become

$$\left(\frac{1}{\xi^2}U^2+V^2+\zeta^2\bar{\phi}^2\right)-\frac{\mu}{c_{11}}\left(U+\zeta(\xi-\xi^2)\bar{\phi}\right)^2\frac{\eta^2+\zeta^2}{\xi^2} = C\left(\frac{1}{\xi^2}U^2+V^2+\zeta^2\bar{\phi}^2\right)^2. \tag{3·16}$$

The second bracket-term on the left-hand side has a coefficient of the second order in ϑ and thus is small compared with the other bracket-term. Thus, as an approximation, the curves of constant intensity may be taken to be

$$\frac{1}{\xi^2}U^2+V^2+\zeta^2\bar{\phi}^2 = \frac{1}{C}. \tag{3·17}$$

These are ellipses having their centres at the extra spot and major and minor axes in the ratio $1:\xi$. Since $\xi \to 1$ the eccentricity is small. The exact curve of constant intensity is one of the fourth degree which touches the ellipse where it is cut by

$$U+\zeta(\xi-\xi^2)\bar{\phi} = 0. \tag{3·18}$$

Since the coefficient of the square of this linear expression is small, the fourth degree curve differs very little from the ellipse considered above. It can be shown to be closed and to lie outside the ellipse and to touch it near the end of the major axis.

4. ANISOTROPIC CASE

To determine the position of the extra spots in the case of an anisotropic substance, even though the degree of anisotropy may be considered as a small perturbation, becomes very complicated, since, in general, equations of the ninth degree are involved. As a simplification only reflexions in the plane of the incident radiation are considered, i.e. directions of scattering for which

$$\eta = 0 \quad \text{and} \quad \bar{\eta} = 0, \tag{4·1}$$

the arrangement of the experiment being as described in the last section. In such a case ξ, ζ and $\bar{\xi}, \bar{\zeta}$ can be expressed in terms of one parameter ϑ and its increment $\bar{\vartheta}$ only, where ϑ is the angle between the incident and the scattered radiation for a Laue spot and $\bar{\vartheta}$ is the angular deflexion from the direction of the Laue spot. The calculation to give $\bar{\vartheta}$ for the extra spot is lengthy but simple. It has been carried out on the assumption that the Laue spots considered are not far removed from the direction of the incident beam. The expressions for $\bar{\vartheta}$ are developed in power series of $\sin\vartheta/2$ or what amounts to the same thing in series of λ/a. It is found that two cases arise. If the solution is written as

$$\bar{\vartheta} = \bar{\phi}(a_0 + a_1 \sin\vartheta/2 + a_2 \sin^2\vartheta/2 + \ldots), \tag{4·2}$$

then the equation for a_0 is $a_0^5 = 0$. The two cases arise according as a_1 is or is not zero. $a_1 = 0$ applies to the following cases: (i) complete isotropy, which is discussed in the last section; or (ii) the indices of the Laue spot concerned have either κ_1 or κ_2 zero; or (iii) κ_1 and κ_3 are equal. In the cases (ii) and (iii) the values of a_2 can be obtained in terms of the elastic constants. They are

$$\left.\begin{aligned} a_2 &= \frac{2c_{11}c_{44}}{(c_{11}-c_{12}-c_{44})(c_{11}+c_{12}+c_{44})}, \quad \text{case (ii)}, \\ a_2 &= \frac{2c_{11}c_{44}-(c_{11}^2-c_{12}^2)}{(3c_{11}+c_{12})(c_{11}-c_{12}-c_{44})}, \quad \text{case (iii)}, \end{aligned}\right\} \tag{4·3}$$

respectively. Taking NaCl and KCl as examples in these cases we find the numerical values of the constants (units of c_{11}, c_{12}, c_{44} do not matter)

	c_{11}	c_{12}	c_{44}	a_2, case (ii)	a_2, case (iii)
NaCl	4770	1294	1320	0·65	0·30
KCl	3750	655	198	0·11	−0·44

In the case of KCl, and more generally for all substances with $c_{11} - c_{12} - 4c_{44} > 0$, the value of a_2 in case (iii) gives a minimum of intensity, and not a maximum.

These solutions are of the same order of magnitude in $\sin\vartheta/2$ as predicted in other theories, but the numerical value of the constant factor is different.

There is not sufficient experimental information about the elastic properties of diamond* to determine the values of a_2. However, the values of these quantities have been determined theoretically by Nagendra Nath (1935). He finds the relation

$$c_{11} \approx c_{12} + c_{44}. \tag{4·4}$$

Now in case (ii) the quantity $(c_{11} - c_{12} - c_{44})$ is a factor in the denominator of a_2. Thus for diamond a_2 becomes very large. If the relation (4·4) is correct one of the roots other than $a_1 = 0$ must be considered. Then the deflexion of the scattered radiation is of the order of $\sin \vartheta/2$. For case (iii) the values of the elastic constants given by Nagendra Nath and the value of a_2 obtained from them are

	c_{11}	c_{12}	c_{44}		
Diamond	9·8	4·5	5·3	$\times 10^{12}$ dyne cm.^{2}	$a_2 = 0·05$.

When none of the conditions (i), (ii) or (iii) is satisfied the equation for a_1, namely,

$$\begin{aligned}
&[a_1(8c_{44} + 4\epsilon \sin^2 2\phi) - 8\epsilon \sin 2\phi \cos 2\phi]\,[c_{11}c_{44}(a_1^2+4)^2 \\
&\qquad\qquad + \epsilon(c_{11}+c_{12})\,(\tfrac{1}{2}a_1^2 \sin 2\phi + 2a_1 \cos 2\phi - 2\sin 2\phi)^2] \\
&\quad - [16c_{11} + 4c_{44}a_1^2 + 4\epsilon \sin 2\phi(\tfrac{1}{2}a_1^2 \sin 2\phi + 2a_1 \cos 2\phi - 2\sin 2\phi)] \\
&\qquad \times [4c_{11}c_{44}a_1(a_1^2+4) + 2\epsilon(c_{11}+c_{12})\,(a_1 \sin 2\phi + 2\cos 2\phi)\,(\tfrac{1}{2}a_1^2 \sin 2\phi \\
&\qquad\qquad + 2a_1 \cos 2\phi - 2\sin 2\phi)] = 0
\end{aligned}$$

has no zero root. I have thought it not worth while to make any numerical calculations from it as there seems to be sufficient evidence obtainable in treating spots with either $\kappa_1 = \kappa_3$ or κ_1 or κ_3 zero.

References

Born, M. 1923 *Atomtheorie des Festen Zustandes.* Leipzig: Teubner.
Born, M. & Sarginson, K. 1941 *Proc. Roy. Soc.* A, **179**, 69.
Debye, P. 1914 *Ann. Phys., Paris*, **43**, 49.
Faxén, H. 1923 *Z. Phys.* **17**, 266.
Jahn, H. A. 1941 *Nature, Lond.*, **147**, 511.
Lonsdale, K. & Smith, H. 1941 *Nature, Lond.*, **148**, 112, 257.
Nagendra Nath, N. S. 1935 *Mem. Indian Inst. Sci.* **1**, 841.
Raman, C. V. & Nilakantan, P. 1940 *Proc. Indian Acad. Sci. Bangalore*, **2A**, 396.
Waller, I. 1925 *Uppsala Univ. Årsskr.*

---

* Mrs Lonsdale & H. Smith (1941) have discovered a new phenomenon for diamond, 'Secondary Spots', which form a small triangle in the neighbourhood of the Laue spot (111), but occur only in the type I crystals. The question whether these spots can be explained with the help of the present theory needs a special investigation.

Cross-section measurements for disintegrations produced by deuterons in the heavy elements

By R. S. Krishnan, Ph.D., 1851 *Exhibition Scholar*
and E. A. Nahum, B.A., *Goldsmiths' Senior Student*
Cavendish Laboratory, Cambridge

(*Communicated by J. D. Cockcroft, F.R.S.—Received* 25 *September* 1941)

The cross-sections have been measured for reactions produced by deuterons in platinum, gold, mercury, thallium, lead, bismuth, thorium and uranium, using deuterons of 9 MeV energy from the Cavendish cyclotron. The observed values of σ for the (d, p) reaction vary from 30×10^{-27} cm.^{2} for platinum to 5×10^{-27} cm.^{2} for uranium. The values for the (d, n) reaction are about one-fifth of those for the (d, p) reaction. Calculations of the cross-sections to be expected on theoretical grounds, have been made for two different values of the standard nuclear radius R_0, namely, $2 \cdot 0 \times 10^{-13}$ and $1 \cdot 47 \times 10^{-13}$ cm. The experimental results are shown to favour the lower of these values.

1. Introduction

We have previously made a general investigation (Krishnan & Nahum 1940, 1941) of the properties of the radio-elements formed by deuteron bombardment of the heavy elements. In this paper an attempt has been made to continue the investigations on more quantitative lines, by measuring the cross-sections for the production by deuteron bombardment of a number of radioactive isotopes of platinum, gold, mercury, thallium, lead, bismuth, thorium and uranium, and by comparing the results with those to be expected on theoretical grounds. So far there has been very little work along these lines. Cork, Halpern & Tatel (1940) have determined the cross-sections for (d, p) and (d, n) reactions in bismuth using deuterons of energy 10 MeV, and have obtained the values 4×10^{-28} and 9×10^{-29} cm.^{2} respectively. A systematic study of cross-sections for the heavy elements promises to be very interesting from the theoretical standpoint, and should give some insight into the mechanism of disintegrations produced by deuterons.

2. Cross-section measurements

The principle of the measurement of cross-section consists in bombarding a very thin layer of the substance under investigation with a known beam of deuterons of definite energy for a given time. From a comparison of the activity resulting from the bombardment with the activity due to a standard source, it is possible to estimate the number of radioactive atoms formed during bombardment. Using the method to be described below, cross-sections for the following reactions have been determined:

(*a*)	(*d*, *p*) reaction	^{196}Pt – ^{197}Pt	(18 hr.)
		^{198}Pt – ^{199}Pt	(32 min.)
		^{197}Au – ^{198}Au	(2·7 days)
		(^{196}Hg – ^{197}Hg)	(48 min.)
		^{204}Hg – ^{205}Hg	(5·5 min.)
		(^{205}Tl – ^{206}Tl)	(4·4 min.)
		^{208}Pb – ^{209}Pb	(2·75 hr.)
		^{209}Bi – ^{210}Bi	(5 days)
		^{232}Th – ^{233}Th	(26·5 min.)
		^{238}U – ^{239}U	(23·5 min.)
(*b*)	(*d*, *n*) reaction	^{197}Au – ^{198*}Hg	(32 hr.)
		(^{200}Hg – ^{201}Tl)	(10·5 hr.)
		^{205}Tl – ^{206*}Pb	(54 hr.)
		^{209}Bi – ^{210}Po	(143 days)

3. EXPERIMENTAL DETAILS

The elements were bombarded in the following form: platinum, gold, thallium, lead and bismuth in the form of thin foils; mercury, thorium and uranium as powders of their compounds. Pure assay foils of platinum and gold were employed, whereas thallium and lead foils were prepared in the laboratory by rolling out pure samples of the metals. Bismuth foils, uniform over a small area, were prepared by evaporating the metal in vacuum on to an aluminium foil. The surface density of the foils used varied from 4 to 12 mg./cm.^{2}. The following technique was adopted for elements bombarded in the form of their compounds. The compounds were finely powdered, and spread as uniformly as possible on weighed copper foils. They were fixed in position by a dilute solution of ciné-film. The copper foils were then cut into pieces of the required shape and the surface density of the powder layer was determined by weighing. Pure oxides of mercury and uranium, and nitrate of thorium were used for bombardment. The uranium oxide was specially purified before bombardment to free it from uranium X.

The foils (or powders) were bombarded by deuterons from the Cavendish cyclotron using a standard target arrangement. The foil was held between a brass cover-plate and the copper target-holder of the cyclotron. The cover-plate was provided with a rectangular aperture of size 1·1 × 0·7 cm., and the size of the foil to be bombarded was always slightly greater than the apertures, so that it filled the whole opening. The target-holder was always fitted in the same position inside the target box. The beam striking the plate was limited by a brass guard-plate provided with an aperture of exactly the same size as that in the cover-plate, and fixed at a distance of 0·3 cm. from the target. The position of this guard-plate was adjusted so that the beam passing through its aperture passed also through the aperture in the cover-plate. This was ascertained by the crude but effective method of observing the scorch mark of the deuteron beam on a piece of paper placed on the target-plate. The target had the usual secondary electron shield. The brass guard-plate was connected to the dee box and was thus insulated from the target. The deflexion of the galvanometer attached to the target-holder was then a true measure of the actual beam striking the foil under investigation.

Bombardments were made under reproducible running conditions. The voltage on the deflecting plate and the radio-frequency voltage on the dees were kept constant as far as possible, the latter being maintained at a high value so as to obtain a well-focused beam. The intensity of the beam striking the target was varied as required by varying the gas pressure or the filament emission. The duration of bombardment, which was mainly determined by the half-life of the activity under investigation, varied from 2 to 30 min., and the beam currents employed varied from $\frac{1}{8}$ to $\frac{1}{2}\,\mu$A. As far as possible a steady value of the beam current was maintained during a bombardment. Readings of the beam current were taken every 10 sec. and the average beam was calculated from these readings.

After each bombardment chemical separations were made. The necessary carriers for the various radioactive isotopes formed were added and the elements were separated in the following form: platinum, gold and mercury as sulphides, thallium as iodide, lead as sulphate, bismuth as hydroxide, thorium as iodate, and uranium as sodium uranyl acetate. The active precipitates were filtered off and spread uniformly on weighed filter papers of standard size (1·8 cm. diameter). The precipitates were washed well and dried. The amount of precipitate collected in each case was determined by weighing and the surface density of the precipitate layer was calculated.

The filter paper carrying the precipitate was placed in a standard specimen holder mounted in a fixed position relative to a standard β-ray counter (which was usually used for β-ray absorption measurements), fitted with a thin mica window through which the β-rays could enter the counter. The activity of the precipitate in this standard position was measured in the usual manner and its decay was followed. The activity of the sample at the end of bombardment was determined by extrapolation of the decay curve. The initial measurement of the activity of the precipitate was followed in every case by a measurement taken with an aluminium absorber interposed between the counter and the substance. In respect of β-particles the absorber used was roughly equivalent in stopping power to the mica window of the counter, the air space between the counter and the source, and a half-thickness of the precipitate layer. It was then possible to correct for losses due to absorption. Where the absorption was small, it was sufficient to assume an exponential law. Where the corrections to be applied were large, it was thought preferable to estimate them by extrapolation of the absorption curves which had been determined previously using the same experimental arrangement. The stopping power of the mica window was equivalent to that of 2 cm. of air, and the specimen holder containing the filter paper was held 2 cm. from the counter window. The thickness of the precipitate layer varied from 2 to 10 mg./cm.^{2}, so that the absorbers used for correction had surface densities varying from 7 to 11 mg./cm.^{2} of aluminium.

Before taking measurements of the activities of the precipitates the counter was standardized using purified uranium oxide. The source used for standardization consisted of 126 mg. of U_3O_8 spread as uniformly as possible over a 1·8 cm. diameter cavity in a standard specimen holder, and fixed in position with a few drops of a solution of ciné-film. This source gave a count of about 1700 per min. when measured in the usual way. Correction for the self-absorption of the standard source and the absorption in the air space and the mica window of the counter was made as follows. A number of similar sources of uranium oxide were prepared with the quantity of uranium oxide varying from 120 to 5 mg. The β-ray activities of these sources were carefully measured under standard conditions and a curve was plotted representing counts per mg. of uranium oxide against the weight of uranium oxide in the source. The curve was extrapolated to zero source thickness. The correction to be applied to the 126 mg. source was found to be 20 %. Taking the decay constant for uranium as $\lambda = 4{\cdot}8 \times 10^{-18}$ sec.$^{-1}$ (Kovarik & Adams 1932; Nier 1939) and calculating on the basis that only one particle per disintegration of uranium is sufficiently energetic to enter the counter, the geometrical efficiency (g) of the counting arrangement was equal to 3·4 %. There was a slight variation from day to day in the efficiency of the counter (less than 2 parts in 100), and this was corrected for in calculating the absolute activity of the precipitate.

From the various measurements, the cross-section for any particular reaction was calculated as follows:

$$\sigma = \frac{N_1}{\beta g \lambda} \frac{1}{n_1 t} \frac{A}{p \rho N},$$

where N_1 = the activity (expressed as counts per sec.) of the sample corrected for absorption, substance lost during chemical separation and extrapolated to end of bombardment,

β = number of electrons emitted per disintegration,

g = geometrical efficiency of the counting arrangement,

λ = disintegration constant in sec.$^{-1}$ of the radioactive isotope formed,

n_1 = the number of deuterons falling on the foil per sec.,

t = duration of bombardment in seconds (small compared with the half-life),

A = atomic weight of the element bombarded,

p = fractional abundance of the isotope responsible for the reaction,

ρ = surface density of the foil under bombardment (g./cm.^{2}),

N = Avogadro's number.

In the case of radioactive bodies with short half-lives the duration of the bombardment (t) was comparable to the half-lives and the necessary corrections were made in the formula for σ.

The cross-section for the formation of polonium from bismuth was not determined directly, but was estimated as follows: A thin layer of bismuth obtained by evaporation on an aluminium foil was bombarded with an intense deuteron beam for an hour. The α-ray activity of the bismuth layer, which was mainly due to polonium formed during bombardment, was measured immediately after the end of bombardment, using an ionization chamber and linear amplifier. The radium E formed during bombardment decayed with a period of 5 days giving rise to polonium. By following the growth and subsequent decay of the α-ray activity of the bombarded bismuth sample, it was possible to estimate the ratio of the number of polonium atoms formed directly to those formed by the decay of Ra E. From a knowledge of the cross-section for Ra E, the cross-section for the formation of polonium could be calculated.

4. RESULTS

The results are summarized in table 1. The values of p, the fractional abundance, are those given by Livingood & Seaborg (1940).

TABLE 1

element	radioactive isotope	period	thickness of foil used mg./cm.^{2}	mean deuteron energy MeV	correction for absorption η	$\beta\sigma \times 10^{+27}$ cm.^{2}	β	$\sigma \times 10^{+27}$ cm.^{2}
${}^{196}_{78}$Pt	${}^{197}_{78}$Pt	18 hr.	10·5	8·93	1·32	15·6 ± 0·9	1	15·6
${}^{198}_{78}$Pt	${}^{199}_{78}$Pt	32 min.	10·5	8·93	1·01	28·8 ± 1·8	1	28·8
${}^{197}_{79}$Au	${}^{198}_{79}$Au	2·7 days	7·3	8·99	1·06	22·0 ± 1·9	1·1?	20·0
(${}^{196}_{80}$Hg	${}^{197}_{80}$Hg)	48 min.	20·0	8·6	1·3	70·0 ± 7·5	?	?
${}^{204}_{80}$Hg	${}^{205}_{80}$Hg	5·5 min.	20·0	8·6	1·03	13·5 ± 0·4	1	13·5
(${}^{205}_{81}$Tl	${}^{206}_{81}$Tl)	4·4 min.	12·0	8·9	1·03	30·3 ± 2·1	1·1?	27·5
${}^{208}_{82}$Pb	${}^{209}_{82}$Pb	2·75 hr.	10·2	8·94	1·2	14·1 ± 0·9	1	14·1
${}^{209}_{83}$Bi	${}^{210}_{83}$Bi	5 days	3·3	9·05	1·03	6·6 ± 0·3	1	6·6†
${}^{232}_{90}$Th	${}^{233}_{90}$Th	26·5 min.	5·0	9·0	1·05	6·3 ± 0·3	1	6·3
${}^{238}_{92}$U	${}^{239}_{92}$U	23·5 min.	8·0	8·98	1·05	5·1 ± 0·5	1	5·1
${}^{197}_{79}$Au	${}^{198*}_{80}$Hg	32 hr.	7·3	8·99	1·7	3·6 ± 0·37	?	?
(${}^{200}_{80}$Hg	${}^{201}_{81}$Tl)	10·5 hr.	20·0	8·6	1·35	1·74 ± 0·18	?	?
${}^{205}_{81}$Tl	${}^{206*}_{82}$Pb	54 hr.	12·0	8·9	1·15	1·05 ± 0·12	?	?
${}^{209}_{83}$Bi	${}^{210}_{84}$Po	143 days	3·3	9·05	—	1·2	1	1·2†

† The values of σ for the formation of ^{210}Bi and ^{210}Po from bismuth obtained by the authors are about 16 times as high as those reported by Cork *et al.* (1940). As they do not specify in their paper how σ was determined, it is not possible to make any further comment on their results.

5. ACCURACY

Although the principle of the method employed for cross-section measurements is a simple one, there are various stages in the experimental procedure where corrections have to be applied, and we discuss below the errors involved.

(*a*) Errors arising from non-uniformity in distribution of the substance under bombardment and consequent variation in surface density were negligible in the case of those elements which were bombarded in the form of foils. The error from this cause would be expected to be appreciable in the case of mercury, thorium and uranium which were bombarded in the form of powders. The deviations observed in the repetitions of the same experiment were in general slightly higher for powders than for foils. But the order of magnitude of these deviations shows that the error introduced by this cause cannot be a serious one. It should be noted that the error introduced by this cause tends to make the observed values of the cross-section smaller than the actual value, since the mean energy of deuterons producing disintegrations is actually less than the mean energy assumed on the basis of uniform distribution.

(*b*) Lack of alignment between the two apertures in the path of the beam, the variation in bending of the beam due to sporadic variations of deflector voltage, and also the process involved in the current integration (described in an earlier section) might be expected to introduce some errors. The following test was made in order to estimate the magnitude of the errors involved. Identical silver foils were bombarded with deuteron beams in the standard arrangement. The beam current and the duration of bombardment were varied for different trials, although the beam current was kept constant as far as possible during each bombardment. The activities of the foils were then measured 7 hr. after the end of bombardment. After reducing the activities to a standard time of bombardment and standard beam current, the maximum deviation observed was 7 %.

(*c*) The presence of a component of very low energy in the deuteron beam will introduce some error. The energy distribution of deuterons in the beam of the cyclotron was experimentally determined using aluminium absorbers, and the results indicate that the low-energy component amounts to 5 % of the measured beam. Since the energy of this component is so low that these deuterons do not produce any appreciable activity in the heavy elements studied, a correction could easily be made for this effect.

(*d*) Weighings were carried out correct to the nearest one-tenth of a milligram, and as the weights of the precipitates involved varied from 5 to 12 mg., the errors involved in weighing should not be greater than ± 2–3 %.

(*e*) In the case of the 32 hr. mercury isotope formed from gold, the chemical separation was incomplete. The small amount of gold present in the mercury fraction was estimated by absorption measurements and a correction applied. In all the other cases the determination of the half-lives indicates that the separations were satisfactory.

(*f*) The various precipitates gave, in the standard counting arrangement, counts varying from 500 to 4000 per min. Counting measurements were taken over a sufficiently long period to make the statistical error involved less than 2 %.

(*g*) When considerable extrapolation and subtraction of activities due to radioactivities of other periods are required, the error involved in estimating the initial activity of a particular radio-element may be as high as 10 %. This occurred only in the case of the 18 hr. ^{197}Pt. In the other cases the error was negligible.

(*h*) In the case of short periods where the β-particles are of high energy, the correction factor (η) for the self-absorption of the β-rays due to the finite thickness of the source and also for the absorption in the air space and in the mica window of the counter was very small and could be accurately estimated. In the case of low-energy β-particles, however, the correction factor was as high as 1·7 (see table 1). In such extreme cases the error involved may be as high as 10 %.

(*i*) The errors arising from non-uniformity in distribution of the precipitate on the filter paper and consequent variation in the geometrical efficiency (g) of the counting system cannot be directly estimated, but the consistency of the results obtained from various measurements suggests that this cause of error cannot be very serious.

(*j*) Differences in the efficiency of counting caused by the possible differences in the back-scattering of the β-rays from the various sources were neglected.

(*k*) The activities produced in the substances by the neutron and γ-ray background of the cyclotron were neglected, except in the case of the 48 min. ^{197}Hg, as experience showed that they were very small compared with those produced by deuterons. In the case of the 48 min. ^{197}Hg the neutron- and γ-ray-induced activity was about 13 % of the total activity. The necessary corrections were made for this effect in the calculation.

The cross-section measurement for the formation of any one radio-element was repeated at least three times. The error limits shown in the table represent the observed maximum deviations from the mean value. The average maximum deviation is 8·5 % (this does not include the error limits for the ^{198*}Hg (32 hr.), where the absorption correction η was very high). If the root mean square had been taken in every case, the average deviation would be very much reduced, and would provide a measure of the *relative* accuracy of the experimental values.

The *absolute* value of the cross-section σ depends on the method of calibration of the counter; the sources of error here are the uncertainty in the decay constant for uranium and the dependence of counter efficiency on the energy of the β-particles.

In all these measurements the quantity that is accurately determined is $\beta\sigma$. Absorption measurements show with a fair degree of certainty that in the case of ^{197}Pt, ^{199}Pt, ^{205}Hg, ^{209}Pb and ^{210}Bi disintegration takes place by the emission of continuous β-rays. In these cases, as there is no quantum radiation present, it can be inferred that β, the number of particles per disintegration, is equal to 1. Consequently the values of σ given in the table should be accurate. In the case of ^{198}Au and ^{206}Tl in addition to continuous β-rays there is a fair proportion of conversion electrons. From a preliminary investigation in a magnetic spectrometer of the particles emitted by these radioactive bodies we have roughly estimated the values of β given in table 1. In the case of ^{197}Hg (48 min.), ^{198*}Hg (32 hr.), ^{201}Tl (10·5 hr.) and ^{206*}Pb (54 hr.) which give conversion electrons only, no accurate estimation of β was possible.

Another factor which may have an important influence on the calculation of σ is the uncertainty in some cases concerning the isotope responsible for the reaction. This point will be considered further in the discussion.

6. Discussion of results

It is generally believed that the formation of the 'compound nucleus' for a (d, n) reaction takes place by the penetration of the potential barrier of the nucleus by the deuteron according to the Gamow-Condon-Gurney process (1928). The (d, p) reaction, however, is believed to occur also by way of another mechanism (known as the Oppenheimer-Phillips process (1935)) in which the deuteron is polarized in the nuclear field and the neutron enters the nucleus.

Let σ_G and $\sigma_{O\cdot P}$ be the cross-sections for the Gamow process and the Oppenheimer-Phillips process respectively.

Then the cross-section for the (d, n) reaction is $\sigma_{G(d,n)}$, while that for the (d, p) reaction

$$= \sigma_{G(d,p)} + \sigma_{O\cdot P}. \qquad (1)$$

The cross-section σ_G for reactions produced by deuterons (Bethe 1937) is given by

$$\sigma_G = \pi R^2 e^{-P} \xi_d (l_c^2/l_0^2)\, \Gamma_a/\Gamma, \qquad (2)$$

where R = the nuclear radius,
e^{-P} = the Gamow penetration probability,
ξ_d = the 'sticking probability',
l_c^2/l_0^2 = the 'angular momentum factor' for the incident deuteron,
Γ_a = the partial probability of emission of the outgoing particle 'a' from the compound nucleus,
Γ = total width of the compound level.

R for any nucleus is usually taken as $R = R_0 A^{\frac{1}{3}}$,

$$P = (4Ze^2/\hbar v)\{\arccos x^{\frac{1}{2}} - x^{\frac{1}{2}}(1-x)^{\frac{1}{2}}\}, \tag{3}$$

with $x = ER/Ze^2$. E is the energy and v the velocity of the incident deuteron.

According to Bethe (1937)

$$l_0 = g = (2MZe^2R)^{\frac{1}{2}}/\hbar \tag{4}$$

and

$$l_c + \tfrac{1}{2} = (\tfrac{1}{2}g)^{\frac{1}{2}}(1-x)^{-\frac{1}{4}}. \tag{5}$$

ξ_d, the 'sticking probability', may be expected to be of the order of unity at the high excitation energies involved. Let us denote $\{\pi R^2 e^{-P}(l_c^2/l_0^2)\}$ by F_G. Then the cross-section σ_G is determined in the main by the factor $F_G\, \Gamma_a/\Gamma$. Thus

$$\sigma_{G(d,n)} = F_G\, \Gamma_n/\Gamma; \quad \sigma_{G(d,p)} = F_G\, \Gamma_p/\Gamma. \tag{6}$$

Various values for R_0 have been suggested, ranging from $2{\cdot}0 \times 10^{-13}$ to $1{\cdot}39 \times 10^{-13}$ cm. The values of F_G have been calculated by the authors for the various elements (see table 2) for 9 MeV deuterons using two different values of R_0, namely, $2{\cdot}0 \times 10^{-13}$ and $1{\cdot}47 \times 10^{-13}$ cm. The value $1{\cdot}47 \times 10^{-13}$ cm. was obtained by Bohr & Wheeler (1939) by computing the nuclear radius in terms of the surface energy of the nucleus.

The cross-section for the O-P process is given by

$$\begin{aligned}\sigma_{\text{O-P}} &= \pi R^2 e^{-P'}(l_c'^2/l_0^2)\,\xi_n\, \Gamma_0/\Gamma \\ &= F_{\text{O-P}}\,\xi_n\, \Gamma_0/\Gamma,\end{aligned} \tag{7}$$

where $e^{-P'}$ = the Oppenheimer-Phillips penetration probability,
ξ_n = the 'sticking probability' for the neutron,
Γ_0/Γ = the probability that no particle is emitted after the neutron is captured.

Calculations by Bethe (1938), Kapur (1937), and Volkoff (1940) show that $\sigma_{\text{O-P}}$ should be greater than σ_G at deuteron energies far below the summit of the potential barrier. Above a certain critical energy (E_c) of the incident deuteron ($E_c = B - I$, where B is the barrier height and I is the dissociation energy of the deuteron), the deuteron enters the nucleus before breaking up, and the Oppenheimer-Phillips process becomes indistinguishable from the Gamow process, and σ for the (d, p) reaction is determined by σ_G. Since Γ_p will be expected to be less than Γ_n, for energies above E_c, $\sigma_{G(d,p)} < \sigma_{G(d,n)}$. This will be the case in our experiments with $R_0 = 2{\cdot}0 \times 10^{-13}$ cm., since E_c is well below 9 MeV for all the nuclei investigated. Therefore in our case $\sigma_{(d,p)} \ll \sigma_{(d,n)}$.

For the lower value of R_0, E_c is above 10·5 MeV for the elements studied. The (d, p) reaction at 9 MeV may therefore be expected to occur mainly by the Oppenheimer-Phillips process. The penetration probability factor, $e^{-P'}$, for this process has been calculated by Volkoff (1940) using $R_0 = 1{\cdot}4 \times 10^{-13}$

cm., and is found to be about 6–7 times the Gamow penetration factor at 9 MeV for these heavy elements. For $R_0 = 1\cdot47 \times 10^{-13}$ cm. this proportionality factor should not be far different from the above value. In our calculations we have provisionally taken the factor as 6. On the assumption that the factor $l_c'^2/l_0^2$ is not far different from l_c^2/l_0^2,† we have calculated $F_{\text{O-P}}$ for the Oppenheimer-Phillips process by multiplying F_G by 6. The results of this calculation are given in column 10 of table 2. It should be noted that since the Oppenheimer-Phillips process amounts to the capture of a neutron, the factor Γ_0/Γ should be large compared with Γ_p/Γ. $F_{\text{O-P}}$ is large compared with F_G at these energies, and the factor Γ_p/Γ tends to reduce still further the importance of σ_G (equation (1)) relative to $\sigma_{\text{O-P}}$. Thus, on the basis of the lower value of R_0, the cross-section $\sigma_{(d,p)}$ at these energies will be determined mainly by $F_{\text{O-P}}$.

TABLE 2

	$R_0 = 2\cdot0 \times 10^{-13}$ cm.				$R_0 = 1\cdot47 \times 10^{-13}$ cm.					Observed	
element	πR^2 $\times 10^{24}$ cm.^{2}	e^{-P} $\times 10$	l_c^2/l_0^2 $\times 10^2$	F_G $\times 10^{27}$ cm.^{2}	πR^2 $\times 10^{24}$ cm.^{2}	e^{-P} $\times 10^2$	l_c^2/l_0^2 $\times 10^2$	F_G $\times 10^{27}$ cm.^{2}	$F_{\text{O-P}}$ $\times 10^{27}$ cm.^{2}	$\sigma_{(d,p)}$ $\times 10^{27}$ cm.^{2}	$\sigma_{(d,n)}$ $\times 10^{27}$ cm.^{2}
$^{195}_{78}$Pt	4·24	7·71	14·0	458	2·29	4·10	6·4	6·0	36	15·8	—
$^{198}_{78}$Pt	4·27	7·91	14·4	485	2·31	4·74	6·4	7·0	42	29·0	—
$^{197}_{79}$Au	4·25	7·11	12·8	387	2·30	3·91	6·3	5·6	34	20·0	~ 3·4
$^{196}_{80}$Hg	4·24	6·52	11·6	321	2·29	3·16	6·1	4·4	26	103·0	—
$^{200}_{80}$Hg	4·30	6·86	12·1	356	2·32	3·43	6·2	4·9	29	—	~ 2·4
$^{204}_{80}$Hg	4·35	7·04	12·6	386	2·35	3·70	6·2	5·4	32	18·0	—
$^{203}_{81}$Tl	4·34	6·47	11·5	322	2·35	3·00	6·1	4·3	26	—	—
$^{205}_{81}$Tl	4·37	6·63	11·7	338	2·36	3·09	6·1	4·4	26	27·5	~ 1·05
$^{208}_{82}$Pb	4·41	6·22	11·1	305	2·38	2·74	6·0	3·9	23	14·1	—
$^{209}_{83}$Bi	4·43	5·73	10·4	264	2·39	2·30	5·9	3·2	19	6·6	1·2
$^{232}_{90}$Th	4·74	3·67	8·2	143	2·56	0·91	5·3	1·25	7·5	6·3	—
$^{238}_{92}$U	4·83	3·11	7·8	116	2·61	0·69	5·2	0·94	5·6	5·1	—

Note. The values of σ given in columns 11 and 12 in the above table are for deuterons of 9 MeV energy. They were calculated from the values given in table 1 and from a knowledge of the corresponding energy-yield curves.

In columns 11 and 12 of table 2 we give the experimental values of σ for the (d, p) and (d, n) reactions respectively (extrapolated to 9 MeV). From a comparative study of the theoretical values of F and the experimental values of σ, the following conclusions may be drawn:

(1) The values of σ determined experimentally for the (d, p) reaction agree fairly closely with the calculated values of $F_{\text{O-P}}$ for $R_0 = 1\cdot47 \times 10^{-13}$ cm., whereas the observed values of σ are one-twentieth to one-fortieth of the

---

† See the argument of Bethe (1938), who shows that the angular momentum factor is of the same order for the two reactions, though slightly higher for the O-P reaction.

theoretical values of $F_{\text{O-P}}\ (= F_G)$ for $R_0 = 2{\cdot}0 \times 10^{-13}$ cm. The agreement between experiment and theory has been achieved with the 4·4 min. thallium isotope by a reassignment of the species responsible. The 4·4 min. thallium was tentatively assigned (Krishnan & Nahum 1940) to ^{204}Tl. On this basis the cross-section for its formation comes to about 70×10^{-27} cm.^{2}, considerably higher than the corresponding values for the other elements of the series studied in these experiments. On the other hand, by reassigning the 4·4 min. period to ^{206}Tl a more reasonable value ($27{\cdot}5 \times 10^{-27}$ cm.^{2}) is obtained. In the same way, the 48 min. ^{197}Hg presents considerable difficulty. On the assumption that it is formed from ^{196}Hg ($p = 0{\cdot}15$ %) by a (d, p) reaction, the minimum value of σ comes to about 70×10^{-27} cm.^{2}. If, however, it is formed from ^{198}Hg ($p = 10{\cdot}1$ %) by a $(d, {}^3\text{H})$ reaction, the minimum value of σ is equal to $1{\cdot}0 \times 10^{-27}$ cm.^{2}. Even so, this value appears to be rather high for a $(d, {}^3\text{H})$ reaction.

(2) In the case of the (d, n) reaction, the experimental value of σ for bismuth (which is more reliable than that for the other elements as it is derived from direct comparison with the cross-section for the (d, p) reaction in that element, and hence does not involve a knowledge of β) agrees well with the value of F_G calculated for the lower value of R_0. The divergence between σ and F_G is very great for $R_0 = 2{\cdot}0 \times 10^{-13}$ cm.

(3) It should be noted that the magnitudes of the observed values of σ for (d, p) and (d, n) reactions considered separately could be explained on the basis of $R_0 = 2{\cdot}0 \times 10^{-13}$ cm. by suitably choosing the values of the factors ξ and Γ_a/Γ. However, when the observed ratio $\sigma_{(d,p)}/\sigma_{(d,n)}$ is considered, it is found to be roughly 5·5 for deuterons of 9 MeV energy. This can only be explained on the basis of the lower value of R_0. For the higher value of R_0 this ratio would be expected to be less than 1, since

$$\sigma_{(d,p)} = F_G\, \xi_d \Gamma_p/\Gamma; \quad \sigma_{(d,n)} = F_G\, \xi_d \Gamma_n/\Gamma$$

and

$$\Gamma_p < \Gamma_n.$$

(4) It has already been shown that the observed energy-yield curves for the (d, n) reaction follow the Gamow curve calculated using the lower value of R_0 (Krishnan 1941; Gant & Krishnan 1941; Krishnan & Nahum 1941). The ratio of the (d, p) yield to the (d, n) yield increases with decreasing energy between 9 and 7·5 MeV. This fact is difficult to explain on the basis of the higher value of R_0, since for that radius $E_c \sim 7{\cdot}5$ MeV.

7. Conclusion

The general trend of the results is clearly in favour of the lower of the alternative values which have been suggested for the standard nuclear radius R_0. Weisskopf & Ewing (1940) have calculated by statistical methods the cross-sections for nuclear reactions produced by protons. They find that the measured cross-sections and their dependence on energy suggest a value of $1{\cdot}3 \times 10^{-13}$ cm. for the standard nuclear radius R_0.

A detailed examination of table 2 shows that although there is general agreement between theory and experiment, there are some irregularities which seem to be outside the range of experimental error; e.g. according to the theory $\sigma_{(d,p)}$ for bismuth should be about three times $\sigma_{(d,p)}$ for thorium, whereas the experimental values are of the same order of magnitude. Certain factors connected with the structure of the nucleus and the nature of the reaction mechanism enter into the exact evaluation of σ, and these are not known with any accuracy at the moment. The simplified theory, however, seems to provide a reasonable explanation of the experimental results.

We are grateful to Professor R. Peierls for some valuable discussions, and to Dr N. Feather for his kind interest in the work.

REFERENCES

Alvarez 1938 *Phys. Rev.* **54**, 486.
Bethe 1937 *Rev. Mod. Phys.* **9**, 69.
Bethe 1938 *Phys. Rev.* **53**, 39.
Bohr & Wheeler 1939 *Phys. Rev.* **56**, 426.
Cork, Halpern & Tatel 1940 *Phys. Rev.* **57**, 371.
Gamow 1928 *Z. Phys.* **51**, 204.
Gant & Krishnan 1941 *Proc. Roy. Soc.* A, **178**, 474.
Gurney & Condon 1928 *Nature, Lond.*, **122**, 439.
Kapur 1937 *Proc. Roy. Soc.* A, **163**, 553.
Kovarik & Adams 1932 *Phys. Rev.* **40**, 718.
Krishnan 1941 *Proc. Camb. Phil. Soc.* **37**, 186.
Krishnan & Nahum 1940 *Proc. Camb. Phil. Soc.* **36**, 490.
Krishnan & Nahum 1941 *Proc. Camb. Phil. Soc.* **37**, 422, and unpublished work.
Livingood & Seaborg 1940 *Rev. Mod. Phys.* **12**, 30.
Nier 1939 *Phys. Rev.* **55**, 150.
Oppenheimer & Phillips 1935 *Phys. Rev.* **48**, 500.
Volkoff 1940 *Phys. Rev.* **57**, 866.
Weisskopf & Ewing 1940 *Phys. Rev.* **57**, 472.

Excitation function measurements for disintegrations produced by deuterons in the heavy elements

By R. S. Krishnan, Ph.D., 1851 *Exhibition Scholar*
and E. A. Nahum, B.A., *Goldsmiths' Senior Student*
Cavendish Laboratory, Cambridge

(*Communicated by J. D. Cockcroft, F.R.S.—Received* 6 *October* 1941)

Excitation function measurements have been made for the (d, p) reaction in thallium, lead, bismuth and thorium, and also for the (d, n) reaction in thallium and bismuth, using deuterons up to 9 MeV energy from the Cavendish cyclotron. The previous results concerning the (d, p) reaction in platinum, gold and uranium and the (d, n) reaction in gold are also included for purposes of comparison. The energy-yield curves obtained are compared with those to be expected on theoretical grounds. The excitation function for the (d, n) reaction is approximately as predicted by the Gamow penetration theory. It is concluded that the Gamow process is responsible for very little if any of the yield of the (d, p) reaction which occurs mainly by the Oppenheimer-Phillips process. The observed ratio of the (d, p) yield to (d, n) yield at 9 MeV and its variation with deuteron energy are shown to favour the lower of the two values of the standard nuclear radius, namely, $1{\cdot}4 \times 10^{-13}$ cm. and $2{\cdot}0 \times 10^{-13}$ cm., which have been considered in this paper.

1. Introduction

The two main reactions which are found to occur in the heavy elements already investigated (Krishnan & Nahum 1940, 1941*a*), when they are bombarded with deuterons up to 9 MeV energy, are the (d, p) and (d, n) reactions. Nuclear reactions in the heavy elements under deuteron bombardment are believed to take place either by formation of a 'compound nucleus' or by the 'Oppenheimer-Phillips process'. According to the compound nucleus picture the incident deuteron is captured and a heavy particle (a neutron, a proton, or an α-particle) is subsequently emitted from the compound nucleus. The yield for this type of reaction is mainly determined by the probability of penetration of the potential barrier of the nucleus by the incident deuteron. This penetration probability is given by the Gamow-Condon-Gurney theory (Gamow 1928; Gurney & Condon 1928) and is very sensitive to the energy of the incident particle. All (d, n) reactions are believed to occur by the above process.

It was found that nuclear reactions of the type ${}^{x}A\ (d, p)\ {}^{x+1}A$ were more probable and less dependent on the energy of the incident particle than would be expected on the Gamow theory. A new mechanism was postulated by Oppenheimer & Phillips (1935) in order to account for these effects. The Oppenheimer-Phillips mechanism may be described as a 'partial entry' of the deuteron into the nucleus. Because of the low dissociation energy of the deuteron, it is split up into a proton and a neutron in the electric field of the nucleus. The proton is repelled and the neutron is captured by the nucleus, leading to a (d, p) reaction. The (d, p) reaction may also be expected to occur by way of the compound nucleus, although the yield will be smaller for low deuteron energies than that for a (d, n) reaction, because of the low probability of emission of the proton from the compound nucleus as compared with the probability of emission of a neutron.

A comparative study of the energy-yield curves for the (d, n) and (d, p) reactions in the heavy elements should provide a useful test as to how far the mechanisms assumed by the Oppenheimer-Phillips and Gamow-Condon-Gurney theories can account for the observed results. Such a study has been made by Hurst, Latham & Lewis (1940), and by Cork, Halpern & Tatel (1940) for bismuth. However, the comparison with theory made by these experimenters is not very conclusive. In the present investigation excitation function measurements have therefore been made for the (d, p) reaction in thallium, lead, bismuth and thorium, and also for the (d, n) reaction in thallium and bismuth. The results obtained by one of us (Krishnan 1941) for the excitation functions for the (d, p) and (d, n) reactions in gold and our previous published results with platinum (Krishnan & Nahum 1941*a*) have been included in this paper for the sake of completeness, together with unpublished results concerning the (d, p) reaction with uranium (Feather & Krishnan).

2. Details of experiment

The method of stacked foils was adopted for determining the excitation functions for the formation of 4·4 min. ^{206}Tl, 2·75 hr. ^{209}Pb and 5-day ^{210}Bi from thallium, lead and bismuth respectively by the (d, p) reaction and 54 hr. ^{206*}Pb and 143-day ^{210}Po from thallium and bismuth respectively by the (d, n) reaction. The method consists in bombarding a pile of foils of the material under investigation with deuterons, and measuring the activity induced in each foil. In this arrangement the energy of the deuterons striking any foil in the pile is determined by the number of foils above it. The foils were cut in the form of a square of side 1·2 cm. The surface density of each foil which was always less than 14 mg./cm.^{2} was determined by weighing. A large number of foils, sufficient to reduce the energy of the 9 MeV deuterons to about 4 MeV, were clamped together on the standard target plate. The foils were pressed together by means of a cover-plate which was bolted on to the target-holder. The cover-plate was provided with an aperture 1 cm. square through which the beam entered. In this arrangement it was ensured that the same beam passed through all the foils and that the region of activation was the same for all the foils. After a suitable bombardment in the cyclotron the foils were stuck on standard specimen holders and the decay of the β-ray activities was followed with the aid of a Geiger counter and a scale-of-eight thyratron recorder. Chemical separation was found to be unnecessary as there was very little interference from other activities. In the case of the 4·4 min. ^{206}Tl a magnetic field was used to resolve the electrons from positrons emitted by the 10·25 min. lead. Decay curves were drawn for the β-ray activity of each foil, and from the family of decay curves for the various foils the energy-yield curve was determined for the formation of the particular radioactive isotope under investigation.

The energy-yield curve for the formation of ^{210}Po from bismuth was determined by counting the α-particles. A pile of bismuth foils (obtained by evaporation on aluminium foils) was given a suitable bombardment, and immediately after bombardment the α-particle activity of each foil was measured by means of an ionization chamber and a linear amplifier. By following the growth and decay of the α-ray activity of the bombarded foils it was possible to estimate the relative number of polonium atoms formed directly in the various foils, and thus to determine the energy-yield curve for polonium.

The excitation function for the formation of the 26·5 min. ^{233}Th from thorium was measured using the 'powder technique'. Targets were prepared by spreading equal quantities (10 mg.) of pure thorium nitrate as uniformly as possible over an area of 1 cm.^{2} on thin copper foils. The layer of powder was fixed in position by the addition of a few drops of a solution of cine-film in amyl acetate. In each case an aluminium foil of 2·5 mg./cm.^{2} covered the deposit. Targets so prepared could be clamped on the target-holder of the cyclotron in a standard position. These targets were bombarded *separately* with deuterons of different specified energies for 5 min. each, keeping the beam steady during each run, but not necessarily at the same value for all the runs. Aluminium foils were used to reduce the energy of the incident deuterons to the required value. After bombardment the thorium nitrate was dissolved in boiling water, and thorium was precipitated as iodate. The thorium fraction was washed well, dried and weighed in order to make correction for any loss of material that had taken place during chemical separation. The β-ray activity was then measured in the usual way.

In order to interpret the results obtained from the above measurements it was necessary to know the initial energy distribution of the deuteron beam. The distribution in range of the particles composing the deuteron beam was determined in the following way. A collector was housed at a distance of 5 cm. from the platinum window in the target plate closing the target chamber of the cyclotron. The ion current to the collector was measured as a function of the superficial mass of the interposed aluminium foil. The results obtained are represented by the curve in figure 1. The conversion of the range in aluminium to the range in air and from that in air to energy was carried out using the data given by Livingston & Bethe (1937). The normal straggling was computed from the formula given by them. The difference between the mean range and the extrapolated range is equal to 2·2 cm. of air, whereas theoretically for an initial energy of 9 MeV it would be 1·5 cm. The energy spread is thus one and a half times that which would be introduced by the straggling of an initially homogeneous beam. After making the necessary corrections for the thickness of the platinum window and also for the bending of the beam, the extrapolated range of the deuterons is found to be equal to 57·4 cm. corresponding to an energy of 9·1 MeV. There is a low-energy component in the beam, well below 10 % in intensity. This low-energy component may possibly arise from scattering of the deuterons by the edges of the grid on which the platinum foil is fixed.

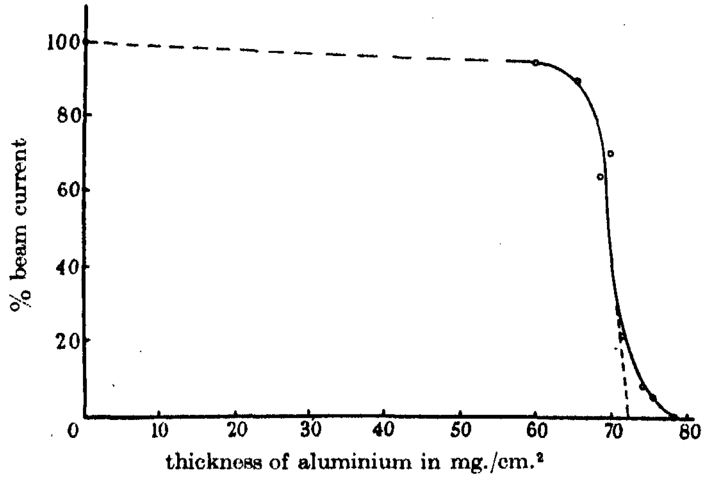


FIGURE 1. Range distribution of deuterons in the beam of the cyclotron.

The finite energy spread of the main beam can be attributed to two main causes: (1) the defect of the experimental method and (2) the energy spread inherent in the cyclotron. The collector was housed in air outside the target chamber and at a finite distance from the window. The correction factor for such a case is difficult to evaluate. The inhomogeneity in the energy of the deuterons due to the cyclotron arises from the finite size of the ion source and the unavoidable variations in the setting of the magnetic field. It should be noted that in the stacked foil method the inhomogeneity in the energy distribution of the deuterons is rapidly increased by the effect of straggling as the beam passes through the foils. The inhomogeneity due to straggling will be of the order of 0·2 MeV when the deuteron energy is reduced to 5 MeV.

3. EXPERIMENTAL RESULTS

The energy-yield curves for the various cases are reproduced in figures 2–6 (curves I and II). In each figure the curves are made to coincide at 9 MeV, and the ordinate scales are arbitrary. The abscissae of points plotted in the energy-yield curves represent the energy in the centre of the foil (or in the centre of the powder layer), of deuterons which have initially the mean energy of the beam, namely, 9·1 MeV. The stopping power of thallium,

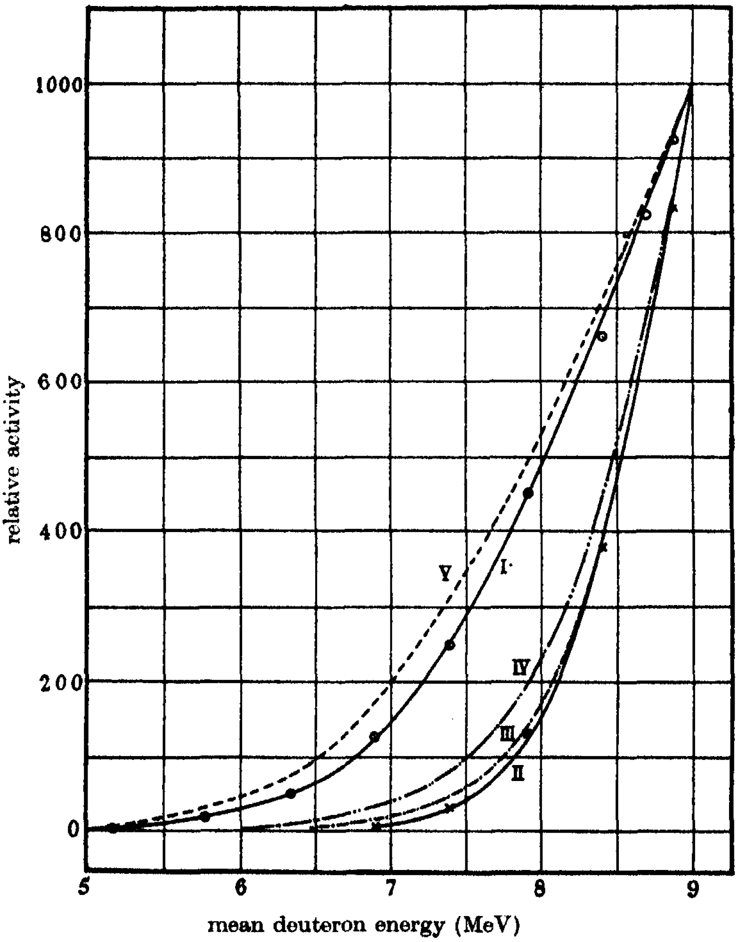
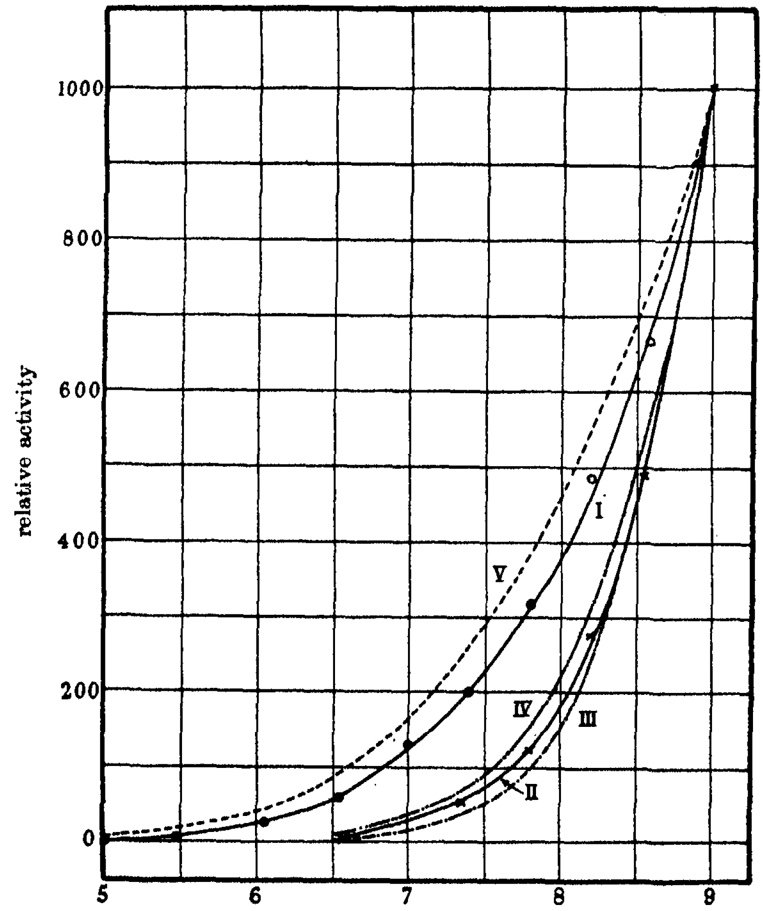


FIGURE 2. Gold.

Curve I, experimental (*d*, *p*) yield 2·7 day ^{198}Au;
II, experimental (*d*, *n*) yield 32 hr. ^{198*}Hg;
III, Gamow curve for $R_0 = 1{\cdot}4 \times 10^{-13}$ cm.;
IV, Gamow curve for $R_0 = 2{\cdot}0 \times 10^{-13}$ cm.;
V, Oppenheimer-Phillips curve for $R_0 = 1{\cdot}4 \times 10^{-13}$ cm.

lead, bismuth and thorium for deuterons was assumed to be the same as that of gold, and the value for gold was taken from the data give by Mano (1934). The activities have been expressed as counts/mg./cm.^{2} of the foil used and no correction was made for the self-absorption of each foil for the electrons it emitted. This does not introduce appreciable error in a single excitation curve since the foils of any particular element varied very little in surface density. In plotting the activities of the foils against the mean energy no correction has been made either for the inhomogeneity of the beam or for the recoil of activated atoms from one foil to those on either side. The necessary corrections have been made for the activities induced by the neutron background of the cyclotron. (This effect was generally less than one-hundredth of the deuteron effect at 9 MeV.)



mean deuteron energy (MeV)

FIGURE 3. Thallium.

Curve I, experimental (*d*, *p*) yield 4·4 min. ^{206}Tl;
II, experimental (*d*, *n*) yield 54 hr. ^{206*}Pb;
III, Gamow curve for $R_0 = 1{\cdot}4 \times 10^{-13}$ cm.;
IV, Gamow curve for $R_0 = 2{\cdot}0 \times 10^{-13}$ cm.;
V, Oppenheimer-Phillips curve for $R_0 = 1{\cdot}4 \times 10^{-13}$ cm.

The duration of bombardment and the intensity of the beam were adjusted in each case so as to give an activity of a few thousand counts per minute in the top foil, and counting measurements were carried out for a sufficiently long time so as to make the statistical error small. Each experiment was repeated at least three times, and the deviations observed in the foil experiments were less than 5 % in the energy region 6·5–9 MeV for the (*d*, *p*) reaction and in the energy region 7·5–9 MeV for the (*d*, *n*) reaction. In the case of thorium the maximum deviation observed was 15 %.

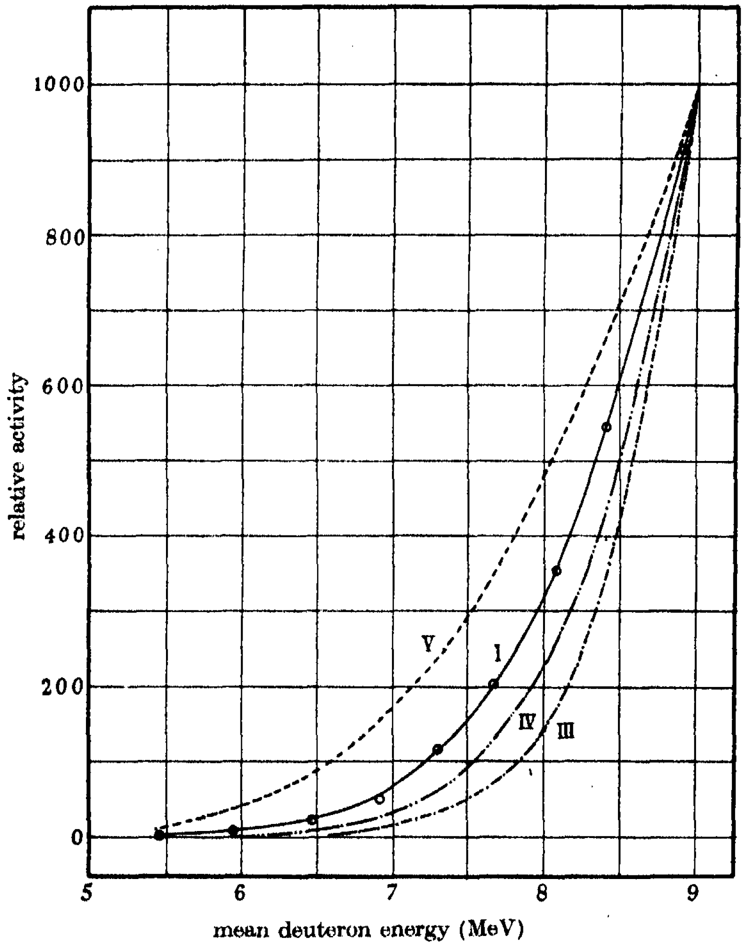


FIGURE 4. Lead.

Curve I, experimental (*d*, *p*) yield 2·75 hr. ^{209}Pb;
III, Gamow curve for $R_0 = 1{\cdot}4 \times 10^{-13}$ cm.;
IV, Gamow curve for $R_0 = 2{\cdot}0 \times 10^{-13}$ cm.;
V, Oppenheimer-Phillips curve for $R_0 = 1{\cdot}4 \times 10^{-13}$ cm.

4. DISCUSSION OF RESULTS

According to Bethe (1937) the cross-section for the (*d*, *n*) reaction taking place by way of a compound nucleus is given by

$$\sigma = \pi R^2 e^{-P}(l_c^2/l_0^2)\,\xi_d(\Gamma_n/\Gamma).$$

In the above formula the 'sticking probability' factor (ξ_d) and the 'partial-width' factor (Γ_n/Γ) can be assumed to be not much dependent on deuteron energy in the energy region 7–9 MeV. The energy dependence of the yield of the reaction will therefore be mainly determined by the variations with energy of the 'Gamow penetrability' factor (e^{-P}) and the 'angular momentum' factor (l_c^2/l_0^2). The values of the factor (e^{-P}) have been calculated for gold, thallium, lead, bismuth and thorium, using Gamow's formula for various deuteron energies and for two different values of the standard nuclear radius R_0, namely, $2{\cdot}0 \times 10^{-13}$ and $1{\cdot}4 \times 10^{-13}$ cm. The values of the factor

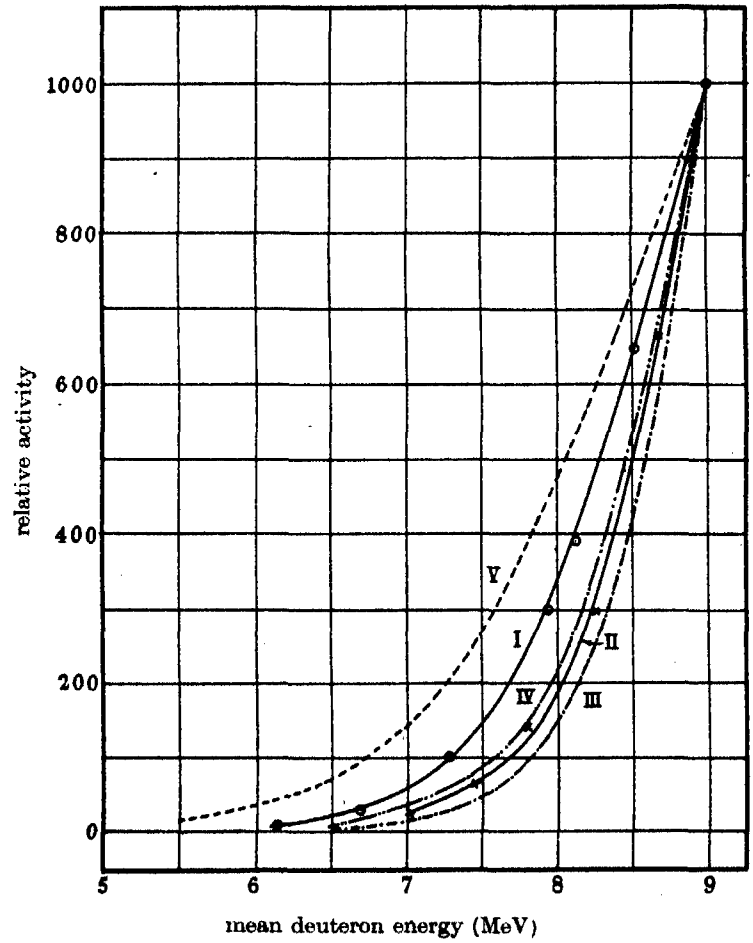


FIGURE 5. Bismuth.
Curve I, experimental (d, p) yield 5-day ^{210}Bi;
II, experimental (d, n) yield 143-day ^{210}Po;
III, Gamow curve for $R_0 = 1{\cdot}4 \times 10^{-13}$ cm.;
IV, Gamow curve for $R_0 = 2{\cdot}0 \times 10^{-13}$ cm.;
V, Oppenheimer-Phillips curve for $R_0 = 1{\cdot}4 \times 10^{-13}$ cm.

(l_c^2/l_0^2) for these cases have been calculated using Bethe's formula (see Bethe 1937, equations (600) and (633)). The variation of the product $(e^{-P}.\,l_c^2/l_0^2)$ with deuteron energy is represented by curve III for $R_0 = 1{\cdot}4\times10^{-13}$ cm. and by curve IV for $R_0 = 2{\cdot}0\times10^{-13}$ cm. in figures 2–6. The theoretical

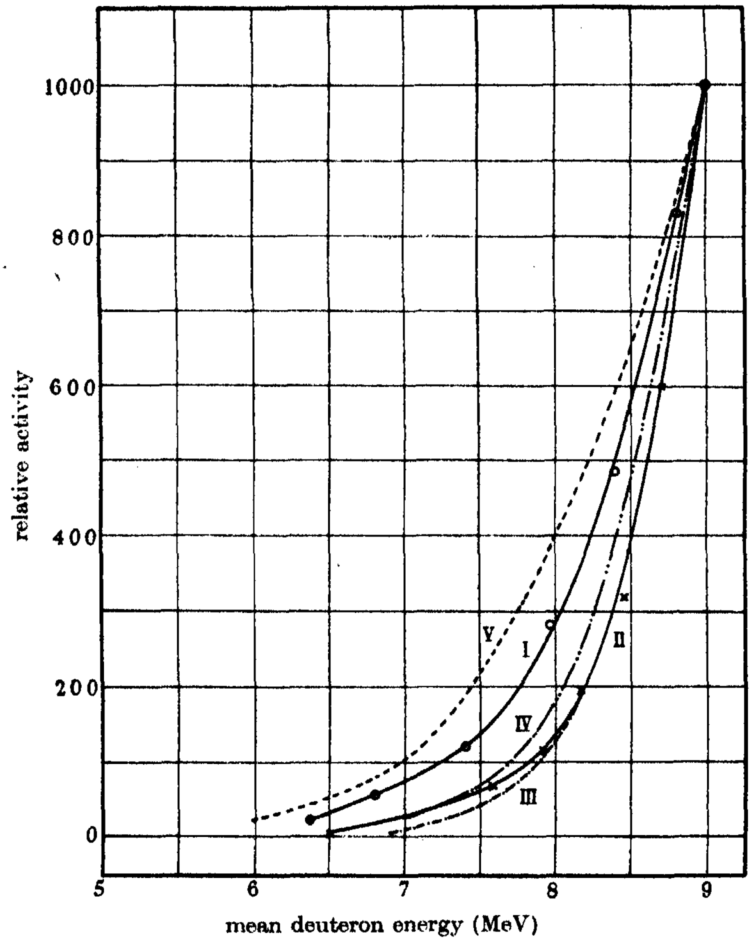


FIGURE 6. Thorium.
Curve I, experimental (d, p) yield 26·5 min. ^{233}Th;
II, experimental fission yield;
III, Gamow curve for $R_0 = 1{\cdot}4\times10^{-13}$ cm.;
IV, Gamow curve for $R_0 = 2{\cdot}0\times10^{-13}$ cm.;
V, Oppenheimer-Phillips curve for $R_0 = 1{\cdot}4\times10^{-13}$ cm.

curves are made to coincide with the experimental energy-yield curves at 9 MeV. It should be remarked that since the factors (ξ_d) and (Γ_n/Γ) are slowly varying factors within the energy region 7–9 MeV, the excitation function for the (d, n) reaction should therefore conform with the curves representing the variation of $(e^{-P}.\,l_c^2/l_0^2)$ with energy as shown in the figures.

If the (d, p) reaction is also assumed to take place according to the Gamow process, the proton emission probability factor enters into the calculation of the excitation function. As this factor increases markedly with energy, the theoretical curve for this reaction should lie on the right of the Gamow curve for the (d, n) reaction, if the curves are made to coincide to 9 MeV.

The cross-section for the Oppenheimer-Phillips process is given by

$$\sigma = \pi R^2 e^{-P'} \xi_n (l_c'^2/l_0^2)\, \Gamma_0/\Gamma.$$

In this case also the factors (ξ_n) and (Γ_0/Γ) may be assumed to vary only slowly with deuteron energy in the energy region 7–9 MeV, and hence the energy dependence of σ will be determined in the main by the variation of $(e^{-P'} . l_c'^2/l_0^2)$. The values of $e^{-P'}$ for the various elements and for the standard nuclear radius $R_0 = 1{\cdot}4 \times 10^{-13}$ cm. have been estimated by extrapolation from the curves showing the ratio $(e^{-P'}/e^{-P})$ reproduced by Volkoff (1940). We have used l_c instead of l_c' because calculations show that the ratio of these two changes very slowly with energy. Curve V in figures 2–6 represents the variation of $(e^{-P'} . l_c^2/l_0^2)$ with energy. The theoretical curve for the Oppenheimer-Phillips process for the higher nuclear radius $(R_0 = 2{\cdot}0 \times 10^{-13}$ cm.) was not plotted as it should coincide with the Gamow curve for deuteron energies greater than the critical deuteron energy E_c. $E_c = (B - I)$, where B is the potential barrier height and I is the dissociation energy of the deuteron. E_c varies from 7·5 MeV for gold to 8·3 MeV for thorium if the higher nuclear radius is assumed.

A comparison of the theoretical curves with the experimental energy-yield curves for the (d, n) reaction shows that the experimental curves lie close to the Gamow curves. We have included in figure 6 the energy-yield curve for the fission of thorium (Gant & Krishnan 1941), as it is known that deuteron-induced fission in thorium takes place by way of a compound nucleus involving Gamow penetration. As there is not much difference in shape between the Gamow curves for the two values of the standard nuclear radius considered here,* the experimental (d, n) curves do not enable us to decide between the two extreme values of the nuclear radius. However, the observed (d, n) curves, except in the case of bismuth, and the fission yield curve for thorium seem to lie closer to the theoretical curve for the lower value of R_0 than to that for the higher value of R_0.

The experimental curves for the (d, p) reaction, on the other hand, lie very much closer to the theoretical curve for the Oppenheimer-Phillips process than to either of the Gamow curves for the (d, n) reaction, and are thus even farther away from the Gamow curve for the (d, p) reaction. It therefore follows that the Gamow process can be responsible for only a very small part, if any, of the yield of the (d, p) reaction, and it must be concluded that the major part of the yield is due to the Oppenheimer-Phillips process. The variation of the ratio of the (d, p) to the (d, n) yield with deuteron energy is given numerically in the following table for gold, thallium, bismuth and thorium. For bismuth this ratio at 9 MeV was experimentally determined as 5 : 1. As it was not easily possible to determine the ratio directly in the other cases, we have arbitrarily taken it to be the same as for bismuth at this energy. This is compatible with the results obtained from cross-section measurements (Krishnan & Nahum 1941 *b*). It will be noticed from the table that the ratio of (d, p) yield to (d, n) yield increases rapidly as the energy is decreased from 9 to 7 MeV. This definitely points to the lower value for the nuclear radius, since on the basis of the higher nuclear radius the Oppenheimer-Phillips process as such would not be expected to take place at these energies. In the case of bismuth, the ratio of (d, p) yield to (d, n) yield at any particular energy appears to be low compared with that for gold and thallium and of the same order as that for thorium. It may be worth pointing out here that a similar discrepancy was observed in cross-section measurements (Krishnan & Nahum 1941 *b*). However, the variation of the $(d, p)/(d, n)$ ratio with energy for bismuth as observed by us is greater than that observed by Hurst *et al.* (1940), while the values of this ratio observed by Cork *et al.* (1940) are of the same order as those reported here for thallium.

---

* Gant & Krishnan (1941) stated that the form of the curve varied considerably with the value of R_0 taken. They had calculated only e^{-P} however, and not $(e^{-P} . l_c^2/l_0^2)$ as we have done here.

TABLE 1. RATIO OF (d, p) YIELD TO (d, n) YIELD

energy in MeV	9	8·5	8	7·5	7
gold	5	8·0	14	28	~70
thallium	5	7·0	11	17	~24
bismuth	5	6·5	9	11	~16
thorium	5	8	11	13	~17

In figure 7 are plotted the energy-yield curves for the (d, p) reaction in platinum, gold, thallium, lead, bismuth, thorium and uranium. They are made to coincide at 9 MeV. The (d, p) curve for platinum coincides with that for gold except for deuteron energies below 6·5 MeV. The results for platinum were taken from our previous work (Krishnan & Nahum 1941 *a*) and those for uranium were taken from the unpublished work of Feather & Krishnan. It will be seen that the curves for the different elements are of similar shape but the ratio of the yield at low energies to that at 9 MeV increases with decreasing atomic number. The only break in this regularity again occurs with bismuth, the curve for which is much nearer to that for thorium than it would normally be expected to be. It is conceivable that the irregular behaviour of bismuth is in some way connected with its high nuclear moment.

The general conclusions concerning the mechanism of deuteron induced reactions and the value of the standard nuclear radius arrived at from these investigations support those deduced from cross-section measurements (Krishnan & Nahum 1942). It is worth noting that the cross-section for some of the reactions investigated here can be estimated at energies lower than 9 MeV by suitably combining the results reported here with our previous measurements of cross-section at 9 MeV.

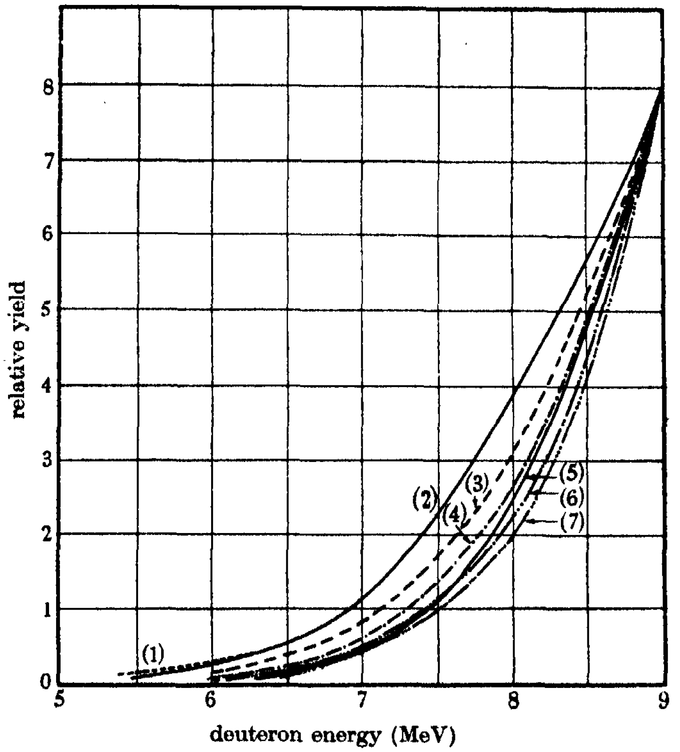


FIGURE 7. Energy yield curves for the (*d*, *p*) reaction.
Curve (1) ------ platinum; (2) ——— gold; (3) ----- thallium;
(4) —·—· lead; (5) ——— bismuth; (6) —··—·· thorium;
(7) —···—···— uranium.

REFERENCES

Bethe 1937 *Rev. Mod. Phys.* **9**, 186–232.
Bethe 1938 *Phys. Rev.* **53**, 39.
Cork, Halpern & Tatel 1940 *Phys. Rev.* **57**, 371.
Feather & Krishnan. Unpublished work.
Gamow 1928 *Z. Phys.* **51**, 204.
Gant & Krishnan 1941 *Proc. Roy. Soc.* A, **178**, 474.
Gurney & Condon 1928 *Nature, Lond.*, **122**, 439.
Hurst, Latham & Lewis 1940 *Proc. Roy. Soc.* A, **174**, 126.
Krishnan 1941 *Proc. Camb. Phil. Soc.* **37**, 186.

Krishnan & Nahum 1940 *Proc. Camb. Phil. Soc.* **36**, 490.
Krishnan & Nahum 1941 *Proc. Camb. Phil. Soc.* **37**, 422.
Krishnan & Nahum 1942 *Proc. Roy. Soc.* A, **180**, 321.
Livingston & Bethe 1937 *Rev. Mod. Phys.* **9**, 245.
Mano 1934 *J. Phys. Radium*, **5**, 626.
Oppenheimer & Phillips 1935 *Phys. Rev.* **48**, 500.
Volkoff 1940 *Phys. Rev.* **57**, 866.

Arrangements with given number of neighbours

By T. S. Chang, Ph.D. and C. C. Ho, B.Sc.

Central University, Chungking, China

(*Communicated by R. H. Fowler, F.R.S.—Received* 29 *September* 1941)

A method previously employed to find the number $p(N, n, X)$ of ways of arranging n **B** particles on N points on a straight line while producing X pairs of **BB** nearest neighbours is extended to find the number of ways of arranging given numbers of **A**, **B**, **C**, ... particles on N points lying in a number of parallel planes containing in each a finite number of points while producing given numbers of various kinds of neighbours (such as the nearest neighbours, the next nearest neighbours, etc.) formed by different pairs of particles. In the former case, the method gives us the limiting value of

$$N^{-1} \log p(N, n, X)$$

as N approaches infinity, and in the present case the limiting value of N^{-1} times the logarithm of our required number as the number of parallel planes approaches infinity. The number of ways of arranging n **B** particles on two rows of points while producing X pairs of **BB** nearest neighbours is investigated in detail, and the results plotted.

Finally, it is pointed out that if we can evaluate certain determinants of infinite number of rows and columns, we can find the number of ways of arrangements when the N points are arranged on a crystalline lattice.

1. INTRODUCTION

Among a group of problems in statistical mechanics collectively described as 'co-operative phenomena', a number can be solved with the standard method of statistical mechanics[†] provided that we know the number $g(N, n, X)$ of ways of arranging $N-n$ A particles and n B particles on a lattice of N points under the condition that X pairs of BB neighbours are formed—or rather the asymptotic value of $g(N, n, X)$ as N, n, X approach infinity, since the N, n, X appearing in the problems are generally very large numbers. At present, the number is known only for the case in which the N points lie on a straight line.[†] If the N points are on lattice structures in which each point is surrounded by the same number of other points, say z, we have an approximate formula[‡] for $\log g(N, n, X)$, which reads

---

[†] I.e. the construction of partition functions or grand partition functions for the assemblies in question. See R. H. Fowler, *Statistical mechanics*, 2nd ed.

$$\begin{aligned}\tfrac{1}{2}zN\log z - \tfrac{1}{2}N(z-2)\log N + (z-1)\,n\log n \\ + (z-1)(N-n)\log(N-n) - X\log 2X - (zn-2X)\log(zn-2X) \\ - (\tfrac{1}{2}Nz - nz + X)\log(Nz - 2nz + 2X) + O(\log N). \qquad (1\cdot1)\end{aligned}$$

Evidently the N points lying on a straight line may be considered to be on one of these lattice structures in which $z = 2$. Thus we may compare the known value of $\log g(N, n, X)$ for this special case with (1·1) after putting $z = 2$, and to our satisfaction we find a perfect agreement.

Owing to the frequent occurrence of the number $g(N, n, X)$ in statistical mechanics it is advisable to investigate it rigorously for several other cases. In the following we develop a method which enables us to find the asymptotic value of $g(N, n, X)$ when the N points lie on a large number of vertical parallel planes containing in each a small number of points. In fact $g(N, n, X)$ for the case with the N points lying on a horizontal straight line was obtained also by this method. In applying it to cases with other arrangements of the N points, the calculations involved are similar, but much more complicated. So we restrict ourselves to the next simplest case in which the N points lie on two horizontal straight lines, or more precisely, lie on the corners of many squares of the same size, whose centres lie on a horizontal straight line and are separated by a distance equal to the sides of the squares.

The method is essentially as follows. If we are interested in $g(N, n, X)$ for a certain arrangement of the N points, let us define a function $\Phi(\xi, \lambda, \eta)$ as the summation of

$$g(N, n, X)\,\xi^N\lambda^n\eta^X \qquad (1\cdot2)$$

over unrestricted values of N, n and X, and proceed to find it. From the definition of Φ and Cauchy's theorem it follows that $g(N, n, X)$ is equal to

$$\frac{1}{(2\pi i)^3}\oint\oint\oint \frac{d\xi\, d\lambda\, d\eta}{\xi^{N+1}\lambda^{n+1}\eta^{X+1}}\Phi(\xi, \lambda, \eta), \qquad (1\cdot3)$$

where the path of integration for each of the variables ξ, λ, η is a circle enclosing the origin taken once in the positive direction in their respective complex planes. After finding $\Phi(\xi, \lambda, \eta)$ and evaluating the integral (1·3) we obtain the required $g(N, n, X)$.

In § 2 we shall describe how to find $\Phi(\xi, \lambda, \eta)$ when the N points lie on a line, and how to carry out the subsequent integration. This has been done in reference 1, but the description given there is incomplete. In § 3 we study the arrangements of any number of kinds of particles on a straight line and introduce a certain function Ψ which is similar to Φ in nature. From this function we shall show in § 5 how to find $\Phi(\xi, \lambda, \eta)$ when the N points lie on two lines. There we shall see that the method is applicable when the N points lie on three or more lines.

[†] Chang, T. S., *Proc. Camb. Phil. Soc.* **35**, 265 (1939), reference 1.

[‡] Reference 1, loc. cit. The formula was first given in this connexion by Ising, *Z. Phys.* **31**, 253 (1925).

2. ARRANGEMENT OF TWO KINDS OF PARTICLES ON A STRAIGHT LINE

Let us denote by $p(A, A, N, n, X)$ the number of ways of arranging $N-n$ A particles and n B particles on N points on a straight line under the restrictions that the first point (i.e. the point on the left end of the line) and the last point (i.e. that on the right end) are occupied by A's and that X pairs of BB neighbours are formed. Let us denote by $p(A, B, N, n, X)$, $p(B, A, N, n, X)$ and $p(B, B, N, n, X)$ similar numbers of ways of arrangements in which the kinds of particles occupying the first and the last points are no longer two A's, but A and B, B and A, and two B's as indicated in the parentheses. Let us consider a certain configuration with A and B particles on a straight line in which the first and the last points are both occupied by A, and let us denote the total number of points in it by N, the number of B's by n and the number of BB neighbours by X. Then, if we introduce a function $\Phi_{p,AA}(\xi, \lambda, \eta)$ and define it as the summation of

$$\xi^N \lambda^n \eta^X \tag{2·1}$$

taken over all such configurations with no restrictions regarding the values of N, n, X, we have immediately $p(A, A, N, n, X)$ given by

$$\frac{1}{(2\pi i)^3} \oint\oint\oint \frac{d\xi\, d\lambda\, d\eta}{\xi^{N+1} \lambda^{n+1} \eta^{X+1}} \Phi_{p,AA}(\xi, \lambda, \eta), \tag{2·2}$$

where the meaning of the integral is obvious. Similarly, we introduce $\Phi_{p,AB}(\xi, \lambda, \eta), \Phi_{p,BA}$ and $\Phi_{p,BB}$, and obtain expressions for $p(A, B, N, n, X), \ldots$ similar to (2·2).

For the construction of $\Phi_{p,AA}$ let us call a chain of $i+1$ points with the first i points all occupied by B and the last point by an A an 'i' chain. Then any configuration with A and B particles on a straight line and with the first and the last points occupied by A's may be considered to consist purely of a number of '0' chains, say m_0+1, a number of '1' chains, say m_1, a number of '2' chains, say m_2, etc., with a '0' chain placed at the left end and the other chains permuted on the straight line in some way or other. Since for configurations starting from the left with a '0' chain, followed by m_0 '0' chains, m_1 '1' chains, etc., permuted in some way or other, the total number of points is always $1+\Sigma_i m_i(i+1)$, the total number of B's always $\Sigma_i m_i i$, and the total number of BB neighbours always $\Sigma'_i m_i(i-1)$, (Σ'_i being different from Σ_i in that the term $i=0$ is excluded), $\Phi_{p,AA}(\xi,\lambda,\eta)$ is nothing but

$$\xi \sum_{m's} \frac{(m_0+m_1+m_2+\dots)!}{m_0!\,m_1!\,m_2!\dots}\,\xi^{\Sigma_i m_i(i+1)}\,\lambda^{\Sigma_i m_i i}\,\eta^{\Sigma'_i m_i(i-1)}, \tag{2·3a}$$

summed over all different values of m_0, m_1, ... For sufficiently small values of ξ, λ, η, (2·3a) reduces by direct summation to

$$\frac{\xi(1-\xi\lambda\eta)}{1-\xi-\xi\lambda\eta-\xi^2\lambda+\xi^2\lambda\eta}, \tag{2·4 a}$$

and hence $p(A,A,N,n,X)$ is equal to

$$\frac{1}{(2\pi i)^3}\oint\oint\oint\frac{d\xi\,d\lambda\,d\eta}{\xi^{N+1}\lambda^{n+1}\eta^{X+1}}\,\frac{\xi(1-\xi\lambda\eta)}{1-\xi-\xi\lambda\eta-\xi^2\lambda+\xi^2\lambda\eta}. \tag{2·5}$$

In a similar way we construct the functions $\Phi_{p,AB}$, $\Phi_{p,BA}$ and $\Phi_{p,BB}$. For example, in constructing $\Phi_{p,AB}$, we observe that any configuration with A and B particles on a straight line and with the first point occupied by an A and the last by a B must be one starting from the left with a '0' chain, followed by a number of '0' chains, a number of '1' chains, etc., permuted in some way and finally followed by a number of points all occupied by B's, and then we have

$$\Phi_{p,AB}(\xi,\lambda,\eta) = \xi\left\{\sum_{m's}\frac{(m_0+m_1+\dots)!}{m_0!\,m_1!\dots}\,\xi^{\Sigma_i m_i(i+1)}\,\lambda^{\Sigma_i m_i i}\,\eta^{\Sigma'_i m_i(i-1)}\right\}\left(\sum_r{}'\,\xi^r\lambda^r\eta^{r-1}\right) \tag{2·3b}$$

summed over all values of m's and r except $r=0$, or

$$\frac{\xi^2\lambda}{1-\xi-\xi\lambda\eta-\xi^2\lambda+\xi^2\lambda\eta}. \tag{2·4b}$$

Similarly, we have

$$\Phi_{p,BA}(\xi,\lambda,\eta) = \Phi_{p,AB}(\xi,\lambda,\eta) = \frac{\xi^2\lambda}{1-\xi-\xi\lambda\eta-\xi^2\lambda+\xi^2\lambda\eta}, \tag{2·4c}$$

$$\Phi_{p,BB}(\xi,\lambda,\eta) = \frac{\xi\lambda(1-\xi)}{1-\xi-\xi\lambda\eta-\xi^2\lambda+\xi^2\lambda\eta}. \tag{2·4d}$$

With $\Phi_{p,AB}$, $\Phi_{p,BA}$ and $\Phi_{p,BB}$ so obtained, we can express $p(A,B,N,n,X)$, $p(B,A)$ and $p(B,B)$ in integrals similar to (2·5).

To integrate (2·5) we write $\Phi_{p,AA}$ as

$$(E-F\lambda)/(G-H\lambda),$$

where E, F, G, H are functions of ξ and η alone. Integration with respect to λ can be easily carried out, giving us

$$p(A,A,N,n,X) = \frac{1}{(2\pi i)^2}\oint\oint\frac{d\xi\,d\eta}{\xi^{N+1}\eta^{X+1}}\left(\frac{H}{G}\right)^n\frac{EH-FG}{GH}. \tag{2·6}$$

For the integration to ξ and η we apply the method of steepest descents, first applied to such problems by Fowler.† The first requirement for the method to be applicable is that N, n, X are large numbers which we shall assume them to be. The second is that H/G must be such a function of ξ and η that when we expand it in powers of ξ and η all the coefficients are positive. This requirement is satisfied since

$$\frac{H}{G} = \frac{\xi\eta + \xi^2 - \xi^2\eta}{1-\xi} = \xi\eta + \frac{\xi^2}{1-\xi}.$$

The third and the last requirement is that the function $(EH - FG)/GH$ must be regular within the path of steepest descent, except possibly with a pole at the origin, and this is also satisfied as can be seen later after determining the saddle point and the path of integration. The saddle point is determined by the set of equations

$$\frac{\partial}{\partial\xi}\left\{n \log\left(\frac{H}{G}\right) - N\log\xi - X\log\eta\right\} = \frac{\partial}{\partial\eta}\left\{n \log\left(\frac{H}{G}\right) - N\log\xi - X\log\eta\right\} = 0, \tag{2·7}$$

which is equivalent to the set

$$\left.\begin{aligned} &G - H\lambda = 0,\\ &\frac{\xi}{N}\left(\frac{\partial(G-H\lambda)}{\partial\xi}\right)_{\lambda,\eta} = \frac{\lambda}{n}\left(\frac{\partial(G-H\lambda)}{\partial\lambda}\right)_{\eta,\xi} = \frac{\eta}{X}\left(\frac{\partial(G-H\lambda)}{\partial\eta}\right)_{\xi,\lambda}.\end{aligned}\right\} \tag{2·8}$$

One of the solutions is

$$\left.\begin{aligned} \xi = \xi_0 = \frac{N-2n+X}{N-n}, \quad \eta = \eta_0 = \frac{X(N-2n+X)}{(n-X)^2},\\ \lambda = \lambda_0 = \frac{N-n}{n}\left(\frac{n-X}{N-2n+X}\right)^2.\end{aligned}\right\} \tag{2·9}$$

Since $n - X$ and $N - 2n + X$ are both positive (lest $p(A, A, N, n, X)$ is zero or pitifully small), ξ_0, λ_0, η_0 are positive and hence the point $(\xi_0, \lambda_0, \eta_0)$ is our saddle point. The quantity $(EH - FG)/GH$ is simply

$$\xi^2/(1-\xi)(\xi + \eta - \xi\eta),$$

which is evidently regular so long as $|\xi| < \xi_0$ and $|\eta| < \eta_0$.

The result of integration therefore gives $p(A, A, N, n, X)$ as

$$\frac{1}{4\pi}\frac{H^n}{\xi^N\eta^X G^n}\frac{EH-FG}{GH}\left\{(-\tfrac{1}{2}n)^3 \frac{D\left(\dfrac{\partial\log G/H}{\partial\log\xi}, \dfrac{\partial\log G/H}{\partial\log\eta}\right)}{D(\log\xi, \log\eta)}\right\}^{-\frac{1}{2}}\left(1 + O\left(\frac{1}{N}\right)\right), \tag{2·10}$$

---

† Refer R. H. Fowler, *Statistical mechanics*, Chapter II.

evaluated at the saddle point. On evaluation it becomes

$$\frac{1}{2\pi}\frac{n^n(N-n)^{N-n}}{X^X(n-X)^{2(n-X)}(N-2n+X)^{N-2n+X}}\left[\frac{N-2n+X}{n(N-n)X}\right]^{\frac{1}{2}}\left(1+O\left(\frac{1}{N}\right)\right). \tag{2·11}$$

In exactly the same way we find $p(A,B)$ and $p(B,B)$ $(p(A,B)=p(B,A))$, and obtain them as (2·11) with the square bracket replaced by

$$\left[\frac{(n-X)^2}{n(N-n)X(N-2n+X)}\right]^{\frac{1}{2}} \quad \text{and} \quad \left[\frac{(n-X)^4}{n(N-n)X(N-2n+X)^3}\right]^{\frac{1}{2}} \tag{2·12}$$

respectively.

Incidentally, we can compute $p(A,A,N,n,X)$, etc., by an entirely different method which we shall describe here for the sake of comparison. Among the $p(A,A,N,n,X)$ ways of arrangements, one part of the ways of arrangement have the second point occupied by an A and the other part the second point occupied by a B. The first part will be $p(A,A,N-1,n,X)$ in number and the second part $p(B,A,N-1,n,X)$. So we have

$$p(A,A,N,n,X) = p(A,A,N-1,n,X)+p(B,A,N-1,n,X), \tag{2·13a}$$

and similarly

$$p(A,B,N,n,X) = p(A,B,N-1,n,X)+p(B,B,N-1,n,X), \tag{2·13b}$$

$$p(B,A,N,n,X) = p(A,A,N-1,n-1,X)+p(B,A,N-1,n-1,X-1), \tag{2·13c}$$

$$p(B,B,N,n,X) = p(A,B,N-1,X)+p(B,B,N-1,n-1,X-1). \tag{2·13d}$$

These equations may be considered as difference equations for the functions $p(A,A)$, $p(A,B)$ and $p(B,B)$ $(p(B,A)=p(A,B))$, and, together with proper initial conditions, may be employed to solve for the functions $p(A,A)$, etc. A solution of (2·13) for $p(A,A)$, $p(A,B)$ and $p(B,B)$ is

$$\frac{(n-1)!\,(N-n-1)!}{X!\,(n-X)!\,(n-X-1)!\,(N-2n+X-1)!}, \tag{2·14a}$$

$$\frac{(n-1)!\,(N-n-1)!}{X!\,[(n-X-1)!]^2\,(N-2n+X)!}, \tag{2·14b, c}$$

$$\frac{(n-1)!\,(N-n-1)!}{X!\,(n-X-1)!\,(n-X-2)!\,(N-2n+X+1)!}. \tag{2·14d}$$

Since these expressions give the correct values of $p(A,A)$, etc., at small values of N, n, X as can be verified directly by counting the number $p(A,A)$, etc., and comparing with (2·14), they are correct expressions for $p(A,A)$, etc., at all N, n and X. It is interesting to see that after replacing the factorials by their Stirling approximations, (2·14) reduce to expressions given by (2·11) and (2·12).

It must not be imagined that this method will help us to get the number of ways of arrangements investigated in §§4 and 5. It is true that we can write out difference equations similar to (2·13). But they cannot be easily solved (or cannot be solved at all) even if we change them into differential equations or do something which in the present instance corresponds to the replacement of $p(A, A, N, n, X)$ by $c_{AA}\xi^N\lambda^n\eta^X$, etc., and the subsequent determination of ξ, λ, η, c_{AA}, etc., both of which we can legitimately do when the number of points is sufficiently large.

3. ARRANGEMENTS OF MORE THAN TWO KINDS OF PARTICLES ON A STRAIGHT LINE

In the discussion of the arrangements of A and B particles on a straight line in the preceding section, A and B do not play symmetrical roles. To preserve the symmetry let n_A, n_B denote the numbers of A and B particles in a certain configuration with A and B particles on a straight line, and $X_{AA}, X_{AB}, X_{BA}, X_{BB}$ the numbers of AA, AB, BA, BB pairs of neighbours. Then instead of $\Phi_{p,AA}$, etc., we introduce a function

$$\Psi'^{(2)}_{AA}(\xi, \lambda_A, \lambda_B, \eta_{AA}, \eta_{AB}, \eta_{BA}, \eta_{BB})\dagger$$

and define it to be the summation of

$$(\xi\lambda_A)^{n_A}(\xi\lambda_B)^{n_B}\eta_{AA}^{X_{AA}}\eta_{AB}^{X_{AB}}\eta_{BA}^{X_{BA}}\eta_{BB}^{X_{BB}}, \tag{3·1}$$

taken over all configurations in which the first and the last points are occupied by A's, with no restrictions regarding the values of the n's and the X's. Similarly, we introduce functions $\Psi'^{(2)}_{AB}$, $\Psi'^{(2)}_{BA}$ and $\Psi'^{(2)}_{BB}$. Then following the procedure of constructing $\Phi_{p,AA}, \ldots$, we construct $\Psi'^{(2)}_{AA}, \ldots$, and obtain

$$\Psi'^{(2)}_{AA} = \frac{\xi\lambda_A(1-\xi\lambda_B\eta_{BB})}{\{(1-\xi\lambda_A\eta_{AA})(1-\xi\lambda_B\eta_{BB})-\xi^2\lambda_A\lambda_B\eta_{AB}\eta_{BA}\}}, \tag{3·2a}$$

$$\Psi'^{(2)}_{AB} = \frac{\xi^2\lambda_A\lambda_B\eta_{AB}}{\{(1-\xi\lambda_A\eta_{AA})(1-\xi\lambda_B\eta_{BB})-\xi^2\lambda_A\lambda_B\eta_{AB}\eta_{BA}\}}, \tag{3·2b}$$

and two other equations for $\Psi'^{(2)}_{BA}$ and $\Psi'^{(2)}_{BB}$ which can be obtained from (3·2*a*) and (3·2*b*) by interchanging the A's and B's.

For the arrangement of three kinds of particles A, B, C on a straight line we shall introduce $\Psi'^{(3)}_{AA}(\xi, \lambda_A, \ldots, \eta_{AA}, \ldots)$ (and similarly $\Psi'^{(3)}_{AB}$, $\Psi'^{(3)}_{AC}$, ..., etc.) defined as the summation of

$$(\xi\lambda_A)^{n_A}(\xi\lambda_B)^{n_B}(\xi\lambda_C)^{n_C}\eta_{AA}^{X_{AA}}\eta_{AB}^{X_{AB}}\ldots\eta_{CC}^{X_{CC}}, \tag{3·3}$$

---

† The number inside the parenthesis on the upper right corner of Ψ' denotes the number of kinds of particles in the arrangement.

taken over all different configurations with A, B, C particles on a straight line in which the first and the last points are occupied by A's, with no restrictions regarding the values of the n's and X's. To construct these $\Psi^{(3)}$'s, let us consider chains of points each of which consists of a number of points occupied by A and B particles and one and only one point occupied by a C particle and located at the right end of the chain. We shall call them μ chains if the first and the last second points are occupied both by A's, ν chains if by A and B, σ chains if by B and A, and τ chains if both by B's. *Of course there are many different kinds of μ chains, different kinds of ν chains, etc., corresponding to different numbers of A and B particles contained in them and different arrangements of these particles.* The different μ chains will be called μ', μ'', ... chains and the numbers of A and B particles and AA, AB, BA, BB neighbours in a μ' chain denoted by $n_{A\mu'}$, $n_{B\mu'}$, $X_{AA\mu'}$, $X_{AB\mu'}$, ..., those in a μ'' chain by $n_{A\mu''}$, $n_{B\mu''}$, $X_{AA\mu''}$, ... Similarly, we have ν', ν'', ... chains, with the numbers $n_{A\nu'}$, $n_{B\nu'}$, $X_{AA\nu'}$, ..., for a ν' chain, etc. Finally, we call a single point occupied by a C particle a 0 chain. To avoid confusion, let us employ the convention that n_A, n_B of none of the μ, ν, σ, τ chains are both zero, i.e. in case n_A, n_B of any of the μ, ν, σ, τ chains are both zero, we shall consider it as a 0 chain. With the various names of the chains defined above, it is easy to see that any configuration with A, B, C particles on a straight line and with C particles at the first and the last points is one starting from the left with a 0 chain, followed by m_0 0 chains, different μ', μ'', ... chains of which we say the numbers are $m_{\mu'}$, $m_{\mu''}$, ..., different ν', ν'', ... chains of which we say the numbers are $m_{\nu'}$, $m_{\nu''}$, ..., etc., arranged in some way on the straight line. So we have $\Psi^{(3)}_{CC}$ equal to

$$\xi\lambda_C \Sigma \frac{(m_0+\Sigma m_{\mu'}+\Sigma m_{\nu'}+\Sigma m_{\sigma'}+\Sigma m_{\tau'})!}{m_0!\,\Pi m_{\mu'}!\,\Pi m_{\nu'}!\,\Pi m_{\sigma'}!\,\Pi m_{\tau'}!}(\xi\lambda_C)^{m_0+\Sigma m_{\mu'}+\Sigma m_{\nu'}+\Sigma m_{\sigma'}+\Sigma m_{\tau'}}\times$$
$$\times\,\eta_{CC}^{m_0}\,\eta_{AC}^{\Sigma m_{\mu'}+\Sigma m_{\sigma'}}\,\eta_{CA}^{\Sigma m_{\mu'}+\Sigma m_{\nu'}}\,\eta_{BC}^{\Sigma m_{\nu'}+\Sigma m_{\tau'}}\,\eta_{CB}^{\Sigma m_{\sigma'}+\Sigma m_{\tau'}}\times$$
$$\times\,\Pi\{(\xi\lambda_A)^{n_{A\mu'}}(\xi\lambda_B)^{n_{B\mu'}}\eta_{AA}^{X_{AA\mu'}}\eta_{AB}^{X_{AB\mu'}}\eta_{BA}^{X_{BA\mu'}}\eta_{BB}^{X_{BB\mu'}}\}^{m_{\mu'}}\times$$
$$\times\,\Pi\{\ldots\}^{m_{\nu'}}\,\Pi\{\ldots\}^{m_{\sigma'}}\,\Pi\{\ldots\}^{m_{\tau'}}, \qquad (3{\cdot}4)$$

where Π denotes product of terms with different numbers of primes, and all summations except the first one denote sums over terms with different number of primes. Simplifying (3·4), we have

$$\xi\lambda_C[1-\{\xi\lambda_C\eta_{CA}\eta_{AC}\Sigma(\xi\lambda_A)^{n_{A\mu'}}(\xi\lambda_B)^{n_{B\mu'}}\eta_{AA}^{X_{AA\mu'}}\eta_{AB}^{X_{AB\mu'}}\eta_{BA}^{X_{BA\mu'}}\eta_{BB}^{X_{BB\mu'}}$$
$$+\,\xi\lambda_C\eta_{CA}\eta_{BC}\Sigma\ldots+\xi\lambda_C\eta_{CB}\eta_{AC}\Sigma\ldots+\xi\lambda_C\eta_{CB}\eta_{BC}\Sigma\ldots+\xi\lambda_C\eta_{CC}\}]. \qquad (3{\cdot}5)$$

The four summations in the above expression are precisely $\Psi^{(2)}_{AA}$, $\Psi^{(2)}_{AB}$, $\Psi^{(2)}_{BA}$ and $\Psi^{(2)}_{BB}$. Substituting their values from (3·2) into (3·5), we get

$$\Psi^{(3)}_{CC} = \xi\lambda_C\{(1-\xi\lambda_A\eta_{AA})(1-\xi\lambda_B\eta_{BB})-\xi^2\lambda_A\lambda_B\eta_{AB}\eta_{BA}\}/D^{(3)}, \qquad (3{\cdot}6a)$$

where $D^{(3)}$ is the determinant

$$\begin{vmatrix} 1-\xi\lambda_A\eta_{AA} & -\xi\lambda_B\eta_{AB} & -\xi\lambda_C\eta_{AC} \\ -\xi\lambda_A\eta_{BA} & 1-\xi\lambda_B\eta_{BB} & -\xi\lambda_C\eta_{BC} \\ -\xi\lambda_A\eta_{CA} & -\xi\lambda_B\eta_{CB} & 1-\xi\lambda_C\eta_{CC} \end{vmatrix}. \tag{3·7}$$

With a slight alternation of (3·4) we find

$$\Psi'^{(3)}_{AC} = \xi^2\lambda_A\lambda_C\{\eta_{AC}+\xi\lambda_B(\eta_{AB}\eta_{BC}-\eta_{BB}\eta_{AC})\}/D^{(3)}. \tag{3·6b}$$

The other $\Psi^{(3)}$ follow from symmetry.

The above shows how to find $\Psi'^{(3)}_{AA}$, $\Psi'^{(3)}_{AB}$, ... from $\Psi'^{(2)}_{AA}$, $\Psi'^{(2)}_{AB}$, ..., and following the same procedure we can find $\Psi'^{(4)}$ from $\Psi'^{(3)}$, $\Psi'^{(5)}$ from $\Psi'^{(4)}$, etc., $\Psi'^{(4)}$, $\Psi'^{(5)}$, ... being defined for the arrangements of four, five, ... kinds of particles in a manner analogous to $\Psi'^{(2)}$ and $\Psi'^{(3)}$. In general, for arrangements of q kinds of particles A, B, ..., Q on a straight line, we have $\Psi'^{(q)}(\xi, \lambda_A, \ldots, \eta_{AA}, \ldots)$ which can be found from $\Psi'^{(q-1)}$. By a mathematical induction the various $\Psi'^{(q)}$'s can be proved to be rational functions of ξ, λ's, η's, of which the denominator is for both $\Psi'^{(q)}_{AA}$ and $\Psi'^{(q)}_{AB}$ the determinant $D^{(q)}$ (symmetrical in *all* the particles A, B, ..., Q)

$$\begin{vmatrix} 1-\xi\lambda_A\eta_{AA} & -\xi\lambda_B\eta_{AB} & -\xi\lambda_C\eta_{AC} & \cdots & -\xi\lambda_Q\eta_{AQ} \\ -\xi\lambda_A\eta_{BA} & 1-\xi\lambda_B\eta_{BB} & -\xi\lambda_C\eta_{BC} & \cdots & -\xi\lambda_Q\eta_{BQ} \\ -\xi\lambda_A\eta_{CA} & -\xi\lambda_B\eta_{CB} & 1-\xi\lambda_C\eta_{CC} & \cdots & -\xi\lambda_Q\eta_{CQ} \\ \vdots & \vdots & \vdots & & \vdots \\ -\xi\lambda_A\eta_{QA} & -\xi\lambda_B\eta_{QB} & -\xi\lambda_C\eta_{QC} & \cdots & 1-\xi\lambda_Q\eta_{QQ} \end{vmatrix}. \tag{3·8}$$

For $\Psi'^{(q)}_{AA}$, the numerator is $\xi\lambda_A$ times the cofactor of $1-\xi\lambda_A\eta_{AA}$ in the determinant $D^{(q)}$, and for $\Psi'^{(q)}_{AB}$, the numerator is $-(\eta_{BA})^{-1}$ times those terms in the determinant $D^{(q)}$ containing η_{BA} as a factor. The mathematical induction is left out and can be easily supplied if desired.

The $\Psi'^{(q)}$ can be used to find the number of ways of arranging q kinds of particles on a straight line with given numbers for various kinds of neighbours. We have in fact the equation

$$p(I, J, n_A, n_B, \ldots, X_{AA}, \ldots) = \frac{1}{(2\pi i)^{q(q+1)}}\oint\cdots\oint\frac{\Pi_S d\lambda_S \Pi_{R,S} d\eta_{RS}}{\Pi_S \lambda_S^{n_S+1}\Pi_{R,S}\eta_{RS}^{X_{RS}+1}}\Phi^{(q)}_{IJ}, \tag{3·9}$$

the notations being natural extensions of those in the preceding. Integration with respect to some of the variables can be easily carried out, and after carrying out integrations with respect to them we find that the right-hand side of (3·9) vanishes except for values of n and X satisfying certain relations. Further integrations can be carried out after the manner of integrating (2·5). The main use of $\Psi'^{(q)}$ is not in this, but in the fact that it helps to determine other numbers of ways of arrangements. The power of $\Psi^{(a)}$ in helping to determine other numbers of ways of arrangements will be illustrated in the following two sections.

4. ARRANGEMENT OF TWO KINDS OF PARTICLES ON A STRAIGHT LINE WITH A LONG-DISTANCE ORDER

In this section we are interested in finding the number of ways of arranging n_α **B**[†] particles on the 1st, 3rd, 5th, ... points of a number of points on a straight line (referred to as α points), η_β **B** particles on the 2nd, 4th, ... points (referred to as β points), with all the points occupied either by **A** or **B** particles, and with the restriction that X pairs of **BB** adjacent neighbours are formed.[‡] We denote the numbers by $f(\mathbf{A}, \mathbf{A}, N, n_\alpha, n_\beta, X)$, $f(\mathbf{A}, \mathbf{B})$, $f(\mathbf{B}, \mathbf{A})$ and $f(\mathbf{B}, \mathbf{B})$, where N denotes the total number of points and the first two letters in the parenthesis the kind of particles in the first and the last points. The calculation for the case in which N is even is slightly different from that for the case in which N is odd, and we proceed now with the first case.

As before we introduce a function $\Phi_{f,\mathbf{AA}}$ and define it to be the summation of

$$\xi^N \lambda_\alpha^{n_\alpha} \lambda_\beta^{n_\beta} \eta^X \tag{4·1}$$

taken over all configurations in which the total number of points is even and the first and the last points are occupied by **A**'s, with no restrictions regarding the values of N, n_α, n_β and X. Now any α point and the β point to its right can be occupied in four different ways, i.e. both by **A**'s, by **A** and **B**, by **B** and **A** and both by **B**'s. Their state of occupation can therefore be represented by figures of the following four types:

A—A	**A—B**	**B—A**	**B—B**
(type A)	(type B)	(type C)	(type D)

Hence if in any configuration with **A** and **B** particles on an even number of points on a straight line we let n_A, n_B, n_C, n_D denote the numbers of the figures of the above four types, X_{AB} (and similarly X_{AC}, X_{AD}, etc.) the number of figures of the type A each of which has a figure of the type B to its right, we see at once that $\Phi_{f,\mathbf{AA}}$ is simply the summation of

$$\xi^{2(n_A+n_B+n_C+n_D)} \lambda_\alpha^{n_C+n_D} \lambda_\beta^{n_B+n_D} \eta^{n_D+X_{BC}+X_{BD}+X_{DC}+X_{DD}}, \tag{4·2}$$

summed over all different configurations with figures of four types arranged on a line, the first figure being of the type A or of the type B, and the last figure either of the type A or C. Therefore $\Phi_{f,\mathbf{AA}}(\xi, \lambda_\alpha, \lambda_\beta, \eta)$ is equal to

---

[†] From now on, we use Clarendon letters to denote the kinds of particles to be arranged, for the letters A, B, C, ... will be employed in another connexion in the same section.

[‡] This has been studied in a separate paper, *Proc. Roy. Soc.* **173**, 48 (1939), where Φ_f is constructed in a different way.

$$\Psi^{(4)}_{AA}(\xi, \lambda_A, \ldots, \eta_{AA}, \ldots) + \Psi^{(4)}_{AC} + \Psi^{(4)}_{BA} + \Psi^{(4)}_{BC} \tag{4·3}$$

after replacing ξ in the $\Psi^{(4)}$'s by ξ^2, λ_A by 1, λ_B by λ_β, λ_C by λ_α, λ_D by $\lambda_\alpha\lambda_\beta\eta$, η_{BC}, η_{BD}, η_{DC}, η_{DD} by η, and all other η's by unity. In this way we obtain

$$\Phi_{f,\mathbf{AA}}(\xi, \lambda_\alpha, \lambda_\beta, \eta) = \xi^2\{1 + \xi^2\lambda_\alpha\lambda_\beta\eta(1-\eta)\}/D_f, \tag{4·4a}$$

where D_f stands for

$$1 - \xi^2(1 + \lambda_\alpha + \lambda_\beta + \lambda_\alpha\lambda_\beta\eta^2) + \xi^4\lambda_\alpha\lambda_\beta(1-\eta)^2.$$

In a similar way we obtain

$$\Phi_{f,\mathbf{AB}}(\xi, \lambda_\alpha, \lambda_\beta, \eta) = \xi^2\lambda_\beta\{1 - \xi^2\lambda_\alpha(1-\eta)\}/D_f, \tag{4·4b}$$

$$\Phi_{f,\mathbf{BA}}(\xi, \lambda_\alpha, \lambda_\beta, \eta) = \xi^2\lambda_\alpha\{1 - \xi^2\lambda_\beta(1-\eta)\}/D_f, \tag{4·4c}$$

$$\Phi_{f,\mathbf{BB}}(\xi, \lambda_\alpha, \lambda_\beta, \eta) = \xi^2\lambda_\alpha\lambda_\beta\{\xi^2 + \eta - \xi^2\eta\}/D_f. \tag{4·4d}$$

To obtain $f(\mathbf{A}, \mathbf{A}, N, n, X)$, it remains to integrate

$$\frac{1}{(2\pi i)^4}\oint\oint\oint\oint \frac{d\xi\, d\lambda_\alpha\, d\lambda_\beta\, d\eta}{\xi^{N+1}\lambda_\alpha^{n_\alpha+1}\lambda_\beta^{n_\beta+1}\eta^{X+1}}\Phi_{f,\mathbf{AA}}(\xi, \lambda_\alpha, \lambda_\beta, \eta). \tag{4·5}$$

As before we write $\Phi_{f,\mathbf{AA}}$ as $(E - F\lambda_\alpha)/(G - H\lambda_\alpha)$ where E, F, G, H are no longer functions of λ_α, integrate with respect to λ_α, and then apply the method of steepest descents, the requirements for the method to be applicable being again all satisfied. The equations determining the saddle point are

$$\frac{\partial}{\partial\xi} = \frac{\partial}{\partial\lambda_\beta} = \frac{\partial}{\partial\eta}\left\{n_\alpha \log\frac{H}{G} - N\log\xi - n_\beta\log\lambda_\beta - X\log\eta\right\} = 0, \tag{4·6}$$

which is equivalent to the set of equations consisting of $D_f \equiv G - H\lambda_\alpha = 0$ and

$$\frac{\xi}{N}\left(\frac{\partial D_f}{\partial\xi}\right)_{\lambda_\alpha,\lambda_\beta,\eta} = \frac{\lambda_\alpha}{n_\alpha}\left(\frac{\partial D_f}{\partial\lambda_\alpha}\right)_{\lambda_\beta,\eta,\xi} = \frac{\lambda_\beta}{n_\beta}\left(\frac{\partial D_f}{\partial\lambda_\beta}\right)_{\eta,\xi,\lambda_\alpha} = \frac{\eta}{X}\left(\frac{\partial D_f}{\partial\eta}\right)_{\xi,\lambda_\alpha,\lambda_\beta}.$$

The saddle point is found to be

$$\left.\begin{aligned} \xi = \xi_0 &= \frac{N - 2n_\alpha - 2n_\beta + X}{(N-2n_\alpha)^{\frac{1}{2}}(N-2n_\beta)^{\frac{1}{2}}}, & \eta = \eta_0 &= \frac{X(N - 2n_\alpha - 2n_\beta + X)}{(2n_\alpha - X)(2n_\beta - X)}, \\ \lambda_\alpha = \lambda_{\alpha0} &= \frac{(N-2n_\alpha)(2n_\alpha - X)^2}{2n_\alpha(N - 2n_\alpha - 2n_\beta + X)^2}, & \lambda_\beta = \lambda_{\beta0} &= \frac{(N-2n_\beta)(2n_\beta - X)^2}{2n_\beta(N - 2n_\alpha - 2n_\beta + X)^2}. \end{aligned}\right\} \tag{4·7}$$

The final expression for $f(\mathbf{A}, \mathbf{A}, N, n, X)$ is

$$\frac{1}{8\pi^{\frac{3}{2}}}\frac{H^{n_\alpha}}{\xi^N\eta^X\lambda_\beta^{n_\beta}G^{n_\alpha}}\frac{EH-FG}{GH}\times$$

$$\times\left\{(-\tfrac{1}{2}n_\alpha)^3\frac{D\left(\dfrac{\partial\log(G/H)}{\partial\log\xi}, \dfrac{\partial\log(G/H)}{\partial\log\lambda_\beta}, \dfrac{\partial\log(G/H)}{\partial\log\eta}\right)}{D(\log\xi, \log\lambda_\beta, \log\eta)}\right\}^{-\frac{1}{2}}\left(1 + O\left(\frac{1}{N}\right)\right), \tag{4·8}$$

evaluated at the saddle point. After some algebra we find $f(\mathbf{A},\mathbf{A})$ as

$$\left[\frac{N-2n_\alpha-2n_\beta+X}{X(2n_\alpha-X)(2n_\beta-X)}\right]^{\frac{1}{2}}\mathbf{L}\left(1+O\left(\frac{1}{N}\right)\right), \tag{4·9}$$

where $\mathbf{L}$ is short for

$$\left(\frac{1}{2\pi}\right)^{\frac{1}{2}}\frac{2^N n_\alpha^{n_\alpha} n_\beta^{n_\beta}(\frac{1}{2}N-n_\alpha)^{\frac{1}{2}N-n_\alpha}(\frac{1}{2}N-n_\beta)^{\frac{1}{2}N-n_\beta}}{X^X(2n_\alpha-X)^{2n_\alpha-X}(2n_\beta-X)^{2n_\beta-X}(N-2n_\alpha-2n_\beta+X)^{N-2n_\alpha-2n_\beta+X}}\times$$
$$\times\left\{\frac{-NX^2+8n_\alpha n_\beta X+4n_\alpha n_\beta(N-2n_\alpha-2n_\beta)}{4n_\alpha n_\beta(N-2n_\alpha)(N-2n_\beta)}\right\}^{\frac{1}{2}}.$$

In a similar way we find $f(\mathbf{AB})$, $f(\mathbf{BA})$ and $f(\mathbf{BB})$ as given by (4·9) with the square bracket replaced by

$$\left[\frac{2n_\beta-X}{X(2n_\alpha-X)(N-2n_\alpha-2n_\beta+X)}\right]^{\frac{1}{2}},\quad\left[\frac{2n_\alpha-X}{X(2n_\beta-X)(N-2n_\alpha-2n_\beta+X)}\right]^{\frac{1}{2}}$$

and

$$\left[\frac{(2n_\alpha-X)(2n_\beta-X)}{X(N-2n_\alpha-2n_\beta+X)^3}\right]^{\frac{1}{2}}$$

respectively.

For the case in which N is odd, we have to introduce a new quantity $\Phi'_{f,\mathbf{AA}}$ which is defined to be the summation of (4·1) over all configurations in which the total number of points is odd and the first and the last points are occupied by $\mathbf{A}$'s. It is not difficult to see that

$$\Phi'_{f,\mathbf{AA}}(\xi,\lambda_\alpha,\lambda_\beta,\eta)=\xi\Phi_{f,\mathbf{AA}}(\xi,\lambda_\alpha,\lambda_\beta,\eta)+\xi\Phi_{f,\mathbf{AB}}, \tag{4·10a}$$

$$\Phi'_{f,\mathbf{AB}}(\xi,\lambda_\alpha,\lambda_\beta,\eta)=\xi\lambda_\alpha\Phi_{f,\mathbf{AA}}(\xi,\lambda_\alpha,\lambda_\beta,\eta)+\xi\lambda_\alpha\eta\Phi_{f,\mathbf{AB}}, \tag{4·10b}$$

apart from unimportant modifications such as the addition of a term linear in ξ, and two other similar equations. These give

$$\Phi'_{f,\mathbf{AA}}(\xi,\lambda_\alpha,\lambda_\beta,\eta)=\xi\{1-\xi^2\lambda_\alpha(1+\lambda_\beta\eta^2)\}/D_f, \tag{4·11a}$$

$$\Phi'_{f,\mathbf{AB}}=\Phi'_{f,\mathbf{BA}}=\xi^3\lambda_\alpha(1+\lambda_\beta\eta)/D_f, \tag{4·11b, c}$$

$$\Phi'_{f,\mathbf{BB}}=\xi\lambda_\alpha\{1-\xi^2(1+\lambda_\beta)\}/D_f. \tag{4·11d}$$

From these expressions similar integrations give $f(\mathbf{A},\mathbf{A},N,n,X)$ as

$$\left[\frac{(N-2n_\alpha-2n_\beta-X)^3}{X(2n_\alpha-X)(2n_\beta-X)(N-2n_\beta)^2}\right]^{\frac{1}{2}}\mathbf{L}'\left(1+O\left(\frac{1}{N}\right)\right), \tag{4·12}$$

where $\mathbf{L}'$ is short for

$$\left(\frac{1}{2\pi}\right)^{\frac{1}{2}}\frac{2^{N+1}n_\alpha^{n_\alpha}n_\beta^{n_\beta}(\frac{1}{2}N+\frac{1}{2}-n_\alpha)^{\frac{1}{2}N+\frac{1}{2}-n_\alpha}(\frac{1}{2}N+\frac{1}{2}-n_\beta)^{\frac{1}{2}N+\frac{1}{2}-n_\beta}}{X^X(2n_\alpha-X)^{2n_\alpha-X}(2n_\beta-X)^{2n_\beta-X}(N+1-2n_\alpha-2n_\beta+X)^{N+1-2n_\alpha-2n_\beta+X}}\times$$
$$\times\left\{\frac{-NX^2+8n_\alpha n_\beta X+4n_\alpha n_\beta(N-2n_\alpha-2n_\beta)}{4n_\alpha n_\beta(N-2n_\alpha)(N-2n_\beta)}\right\}^{\frac{1}{2}}.$$

and $f(\mathbf{AB})$ and $f(\mathbf{BB})$ ($f(\mathbf{AB}) = f(\mathbf{BA})$) as (4·12) with the square bracket replaced by

$$\left[\frac{(2n_\alpha - X)(N - 2n_\alpha - 2n_\beta + X)}{X(2n_\beta - X)(N - 2n_\beta)^2}\right]^{\frac{1}{2}}$$

and

$$\left[\frac{(2n_\alpha - X)^3}{X(2n_\beta - X)(N - 2n_\alpha - 2n_\beta + X)(N - 2n_\beta)^2}\right]^{\frac{1}{2}}$$

respectively. It can be noticed that the various f's cannot be simply products of factorials as the p's.

An extension is to find the number of ways of arranging n_α **B** particles on the 1st, 4th, 7th, ... points (α points), n_β **B** particles on the 2nd, 5th, ... points (β points), n_γ **B** particles on the 3rd, 6th, ... points (γ points), with points either occupied by **A** or **B**, and with the restriction that X pairs of **BB** neighbours are formed and some other restrictions U such as those regarding the state of occupation of the first and the last points (or some other points) or regarding the value of the remainder obtained after dividing the total number of points by three. To study the number we introduce a function Φ_f^* defined to be the summation of

$$\xi^N \lambda_\alpha^{n_\alpha} \lambda_\beta^{n_\beta} \lambda_\gamma^{n_\gamma} \eta^X$$

over all configurations satisfying the restrictions U. It is not difficult to see that Φ_f^* can be obtained from the different $\Psi^{(8)}$ or their combinations by a replacement of the variables, and hence that for different restrictions U, the denominator D_f^* of Φ_f^* is always the same. In fact D_f^* is

$$1 - \xi^3\{1 + \lambda_\alpha + \lambda_\beta + \lambda_\gamma + \eta(\lambda_\alpha\lambda_\beta + \lambda_\beta\lambda_\gamma + \lambda_\gamma\lambda_\alpha) + \eta^3\lambda_\alpha\lambda_\beta\lambda_\gamma\} +$$
$$+ \xi^6\lambda_\alpha\lambda_\beta\lambda_\gamma(\eta - 1)^3.$$

After obtaining Φ_f^*, the next thing is to perform an integration which can be carried out exactly as before. In so doing the logarithm of the integral is found to fall into two parts, one of the order $\log N$ or bigger, and the other of the order unity or smaller. The first part depends upon the denominator and is therefore the same for different U. The second part depends upon the numerator and is therefore different for different U. If we want the first part only, which is by far the most important part, we shall be solely interested in the denominator. The final result for the logarithm of the number is

$$-N\log\xi - n_\alpha\log\lambda_\alpha - n_\beta\log\lambda_\beta - n_\gamma\log\lambda_\gamma - X\log\eta - 2\log N + O(1), \quad (4{\cdot}13)$$

ξ, λ_α, λ_β, λ_γ and η being the unique positive root of $D_f^* = 0$ and

$$\frac{\xi}{N}\frac{\partial D_f^*}{\partial\xi} = \frac{\lambda_\alpha}{n_\alpha}\frac{\partial D_f^*}{\partial\lambda_\alpha} = \frac{\lambda_\beta}{n_\beta}\frac{\partial D_f^*}{\partial\lambda_\beta} = \frac{\lambda_\gamma}{n_\gamma}\frac{\partial D_f^*}{\partial\lambda_\gamma} = \frac{\eta}{X}\frac{\partial D_f^*}{\partial\eta}.$$

So far we are only interested in knowing the number of ways of arrangements in which the number of pairs of adjacent neighbours formed by **B** particles is equal to a given number. We can also require the number of pairs of next nearest neighbours formed by the **B**'s to be equal to another given number. For definiteness we say that we want the part of $f(\ldots, N, n, X)$ ways of arrangements in which Y pairs of next nearest neighbours† are formed by **B**'s. Then we introduce a function Φ_f^{**} defined as the summation of

$$\xi^N \lambda_\alpha^{n_\alpha} \lambda_\beta^{n_\beta} \eta^X \zeta^Y$$

over all configurations. The reasoning employed in obtaining Φ_f from $\Psi^{(4)}$ can be taken over here, and it is easy to see that the denominator D_f^{**} of Φ_f^{**} is simply $D^{(4)}(\xi, \lambda_A, \ldots, \eta_{AA}, \ldots)$ with ξ replaced by ξ^2, λ_A by 1, λ_B by λ_β, λ_C by λ_α, λ_D by $\lambda_\alpha \lambda_\beta \eta$, η_{BB}, η_{CC}, η_{CD}, η_{DB} by ζ, η_{BD}, η_{DC} by $\eta\zeta$, η_{BC} by η, η_{DD} by $\eta\zeta^2$, and all other η's by unity. The denominator D_f^{**} is therefore equal to

$$\begin{aligned}1 - \xi^2\{1 + \zeta(\lambda_\alpha + \lambda_\beta) + \lambda_\alpha \lambda_\beta \eta^2 \zeta^2\} + \xi^4\{(\lambda_\alpha + \lambda_\beta)(\zeta - 1)(1 + \lambda_\alpha \lambda_\beta \eta^2 \zeta^2) + \\ + \lambda_\alpha \lambda_\beta (\zeta^2 + \eta^2 \zeta^2 - 2\eta)\} - \xi^6 \lambda_\alpha \lambda_\beta (\zeta - 1)^2 \{1 + (\lambda_\alpha + \lambda_\beta)\eta^2 \zeta + \lambda_\alpha \lambda_\beta \eta^2 \zeta^2\} + \\ + \xi^8 \lambda_\alpha^2 \lambda_\beta^2 \eta^2 (\zeta - 1)^2.\end{aligned}$$

The required integration is, however, slightly different from those described above. It is extremely similar to the integration described in the following section, and will therefore not be repeated here. It turns out that, as before, the numerator is immaterial to the expression for the logarithm of the number so long as we neglect terms of the order unity. Loosely speaking, the logarithm of the number is

$$-N \log \xi - n_\alpha \log \lambda_\alpha - n_\beta \log \lambda_\beta - X \log \eta - Y \log \zeta - 2 \log N + O(1), \quad (4{\cdot}14)$$

ξ, λ_α, λ_β, η and ζ being the positive root of $D_f^{**} = 0$ and

$$\frac{\xi}{N}\frac{\partial D_f^{**}}{\partial \xi} = \frac{\lambda_\alpha}{n_\alpha}\frac{\partial D_f^{**}}{\partial \lambda_\alpha} = \frac{\lambda_\beta}{n_\beta}\frac{\partial D_f^{**}}{\partial \lambda_\beta} = \frac{\eta}{X}\frac{\partial D_f^{**}}{\partial \eta} = \frac{\zeta}{Y}\frac{\partial D_f^{**}}{\partial \zeta},$$

giving the largest value of (4·14).

5. ARRANGEMENT OF TWO KINDS OF PARTICLES ON TWO STRAIGHT LINES

Now we come to find the number $h(N, n, X)$‡ of ways of arranging $N - n$ **A** particles and n **B** particles on N points lying on the corners of many squares of the same size, whose centres lie on a horizontal straight line and are separated by a distance equal to the sides of the squares. For simplicity we refer to these N points as two rows of $\frac{1}{2}N$ points each.

---

† In the present instance they are points or particles separated by a third point lying between them.

‡ We employ this new notation to avoid confusion with $g(N, n, X)$ which is the general notation for the number of ways of arrangements of **A** and **B** particles on N points, of which the arrangement is unspecified.

Our procedure of finding $h(N, n, X)$ is essentially the same as before. Let the total number of points in a certain configuration with **A** and **B** particles on two rows of points as described above be N, the number of **B**'s n, and the number of pairs of **BB** neighbours X. Then we introduce a function $\Phi_h(\xi, \lambda, \eta)$ defined as the summation of

$$\xi^N \lambda^n \eta^X \tag{5·1}$$

over all such configurations with no restrictions regarding the values of N, n and X, and have $h(N, n, X)$ given by

$$\frac{1}{(2\pi i)^3} \oint\oint\oint \frac{d\xi\, d\lambda\, d\eta}{\xi^{N+1} \lambda^{n+1} \eta^{X+1}} \Phi_h(\xi, \lambda, \eta). \tag{5·2}$$

In any configuration with **A** and **B** particles on two rows of points, any point on the top row and the point directly below it may be both occupied by **A**'s, or by **A** and **B**, or by **B** and **A**, or both by **B**'s. The different states of occupation can be accordingly represented by figures of the following types:

$$\begin{array}{cccc} \mathbf{A} & \mathbf{A} & \mathbf{B} & \mathbf{B} \\ | & | & | & | \\ \mathbf{A} & \mathbf{B} & \mathbf{A} & \mathbf{B} \\ (\text{type } A) & (\text{type } B) & (\text{type } C) & (\text{type } D) \end{array}$$

If in a certain configuration we let n_A, n_B, n_C, n_D denote the numbers of figures of the above four types, X_{AB} (and similarly X_{AC}, X_{AD}, etc.) the number of figures of the type A each of which has a figure of the type B to its right, we see at once that $\Phi_h(\xi, \lambda, \eta)$ is simply the summation of

$$\xi^{2(n_A+n_B+n_C+n_D)} \lambda^{n_B+n_C+2n_D} \eta^{n_D+X_{BB}+X_{CC}+X_{BD}+X_{CD}+X_{DB}+X_{DC}+2X_{DD}} \tag{5·3}$$

taken over all configurations with the four kinds of figures arranged on a straight line, with no restrictions regarding the values of n's and X's. Hence $\Phi_h(\xi, \lambda, \eta)$ can be obtained from $\Psi^{(4)}(\xi, \lambda_A, \ldots, \eta_{AA}, \ldots)$ by replacing ξ in $\Psi^{(4)}$ by ξ^2, λ_A by 1, λ_B, λ_C by λ, λ_D by $\lambda^2\eta$, X_{BB}, X_{CC}, X_{BD}, X_{CD}, X_{DB}, X_{DC} by η, X_{DD} by η^2, and all other η's by unity. In this way we obtain the denominator D_h of Φ_h as

$$\begin{aligned} 1 - \xi^2(1 + 2\lambda\eta + \lambda^2\eta^3) + \xi^4\lambda(\eta - 1)\{2 + \lambda(\eta + 1)^2 + 2\lambda^2\eta^3\} - \\ - \xi^6(\eta - 1)^2 \lambda^2(1 + 2\lambda\eta^2 + \lambda^2\eta^3) + \xi^8\lambda^4\eta(\eta - 1)^4, \end{aligned}$$

or
$$\frac{1-a}{\lambda(\eta - 1)}\{-a(1-a) + \lambda b - \lambda^2\eta^3 a(1-a)\}, \tag{5·4}$$

where a is short for $\xi^2\lambda(\eta - 1)$ and b for

$$-(1+a) + \eta(1 - a + a^2 + a^3) + \eta^2 a^2(1-a).$$

As can be seen below, the numerator of Φ_h is unimportant in the determination of $h(N, n, X)$ so long as we do not want terms of the order unity in the final expression for the logarithm of $h(N, n, X)$, and hence it will not be given.

Now we are faced with the task of integrating (5·2). Let us first transform the variables of integration ξ, λ, η to ν, λ, η, where $\nu = \xi^2\lambda$. If we put $D_h = 0$ and solve for λ in terms of ν and η, we obtain

$$\lambda = \lambda_1(\nu, \eta) = \{b + [b^2 - 4\eta^3 a^2(1-a)^2]^{\frac{1}{2}}\}/2\eta^3 a(1-a), \tag{5·5a}$$

and
$$\lambda = \lambda_2(\nu, \eta) = \{b - [b^2 - 4\eta^3 a^2(1-a)^2]^{\frac{1}{2}}\}/2\eta^3 a(1-a). \tag{5·5b}$$

Splitting Φ_h into partial fractions, say

$$C_1(\nu, \eta)/\{\lambda - \lambda_1(\nu, \eta)\}, \quad C_2(\nu, \eta)/\{\lambda - \lambda_2(\nu, \eta)\}, \tag{5·6}$$

and a polynomial in λ and integrating (5·2) with respect to λ, we get

$$h(N, n, X) = I_1 + I_2, \tag{5·7}$$

where
$$I_{1,2} = \frac{1}{(2\pi i)^2} \oint\oint \frac{d\nu\, d\eta}{\nu^{\frac{1}{2}N+1}\eta^{X+1}} C_{1,2}(\nu, \eta)\, [\lambda_{1,2}(\nu, \eta)]^{\frac{1}{2}N-n}. \tag{5·8}$$

Now we apply the method of steepest descent to I_1 and I_2. For clarity we restrict ourselves for the moment to I_1.

We start, as usual, with the determination of the saddle point by solving the set of equations

$$\frac{\partial}{\partial\nu} = \frac{\partial}{\partial\eta} \{(\tfrac{1}{2}N - n)\log\lambda_1(\nu, \eta) - \tfrac{1}{2}N\log\nu - X\log\eta\} = 0, \tag{5·9}$$

which is equivalent to the set of equations

$$\frac{1}{\frac{1}{2}N}\left[\nu\left(\frac{\partial D_h}{\partial\nu}\right)_{\lambda,\eta}\right]_{(\lambda_1)} = \frac{1}{n - \frac{1}{2}N}\left[\lambda\left(\frac{\partial D_h}{\partial\lambda}\right)_{\eta,\nu}\right]_{(\lambda_1)} = \frac{1}{X}\left[\eta\left(\frac{\partial D_h}{\partial\eta}\right)_{\nu,\lambda}\right]_{(\lambda_1)}, \tag{5·10}$$

where the suffix (λ_1) is introduced to indicate the fact that the functions with the suffix are evaluated at $\lambda = \lambda_1(\nu, \eta)$. Owing to the fact that the coefficients a_{lm} in the expansion $\Sigma a_{lm}\nu^l\eta^m$ of the function $\lambda_1(\nu, \eta)$ in powers of ν and η are not all positive, the equations (5·9) or (5·10) may admit several sets of positive real solutions (or none) inside the domain of convergence of the series $\Sigma a_{lm}\nu^l\eta^m$. Among the various sets of solutions, real or otherwise, we shall choose the set, say (ν_s, η_s), which corresponds to a minimum of

$$(\tfrac{1}{2}N - n)\log\lambda_1(\nu, \eta) - \tfrac{1}{2}N\log\nu - X\log\eta, \tag{5·11}$$

if we confine ν to be on the line in ν-plane joining ν_s to the origin and η on the line in η-plane joining η_s to the origin, and therefore to a maximum of (5·11) if we confine ν and η to be on the circles $|\nu| = |\nu_s|$ and $|\eta| = |\eta_s|$. Suppose that for suitably given values of N, n, X there exists one such set of solutions ν_s and η_s. Then on letting the point be our saddle point, i.e. by choosing the path of integration to be circles $|\nu| = |\nu_s|$, $|\eta| = |\eta_s|$, we can apply the method of steepest descents as usual and obtain $\log I_1$ as

$$(\tfrac{1}{2}N - n)\log\lambda_1(\nu_s, \eta_s) - \tfrac{1}{2}N\log\nu_s - X\log\eta_s - \log N + O(1). \tag{5·12}$$

Evidently there cannot be another set of solutions with the same nature, for if there is, say (ν_s', η_s'), we may let (ν_s', η_s') be the saddle point and obtain $\log I_1$ as (5·12) with ν_s' and η_s' in place of ν_s and η_s, which is clearly impossible, since $\log I_1$ must not take two different values in two different ways of evaluation. Further, ν_s, η_s must be real and positive, for, as can be seen later, $h(N, n, X)$ is practically either I_1 or I_2. If it is I_2 we can leave off our investigation of the present ν_s and η_s altogether. If it is I_1, then I_1 must be real and positive. This, combined with

$$\left.\begin{aligned} \frac{I_1(N+1, n, X)}{I_1(N, n, X)} = \left(\frac{\lambda_1(\nu_s, \eta_s)}{\nu_s}\right)^{\frac{1}{2}}, \quad \frac{I_1(N, n+1, X)}{I_1(N, n, X)} = \frac{1}{\lambda_1(\nu_s, \eta_s)}, \\ \frac{I(N, n, X+1)}{I(N, n, X)} = \frac{1}{\eta_s}, \qquad\qquad \end{aligned}\right\} \tag{5·13}$$

requires ν_s, η_s, $\lambda_1(\nu_s, \eta_s)$ to be real and positive. It will be seen to be actually the case.

For the particular case $n = \tfrac{1}{2}N$, the saddle point of I_1 can be determined as follows. At $n = \tfrac{1}{2}N$, equation (5·10) requires

$$\left[\left(\frac{\partial D_h}{\partial\lambda}\right)_{\eta,\nu}\right]_{(\lambda_1)} = 0, \quad \text{or} \quad b - 2\eta^3 a(1-a)\lambda_1 = 0. \tag{5·14}$$

Combining with $D_h(\lambda_1) = 0$ we have $\lambda_1 = \pm\eta^{-\frac{1}{2}}$. With $\lambda_1 = +\eta^{-\frac{1}{2}}$, equation (5·14) gives $\nu = \eta^{-\frac{1}{2}}(\eta - 1)^{-1}$ or

$$\nu = \frac{-(\eta+1) \pm \{(\eta+1)^2 - 4\eta^{\frac{1}{2}}(1-\eta^{\frac{1}{2}})^2\}^{\frac{1}{2}}}{2\eta^{\frac{1}{2}}(1-\eta^{\frac{1}{2}})(\eta-1)}. \tag{5·15}$$

Combining the equation (5·15) with + sign before the square root and

$$\frac{1}{\tfrac{1}{2}N}\left(\nu\frac{\partial D_h}{\partial\nu}\right)_{(\lambda_1)} = \frac{1}{X}\left(\eta\frac{\partial D_h}{\partial\eta}\right)_{(\lambda_1)}, \tag{5·10}$$

we get
$$\frac{2X}{N} = \frac{1}{4(\eta-1)}\left\{\frac{(\eta^{\frac{1}{2}}+1)^2(\eta - 4\eta^{\frac{1}{2}} + 1)}{[(\eta+1)^2 - 4\eta^{\frac{1}{2}}(1-\eta^{\frac{1}{2}})^2]^{\frac{1}{2}}} + 5\eta - 1\right\}. \tag{5·16}$$

It can be verified that for suitably given values† of $2X/N$, (5·16) has only one positive solution for η, say η_0, that ν_0 obtained by substituting η_0 into the right-hand side of (5·15) with + sign before the square root is also positive, and that (ν_0, η_0) corresponds to a minimum of (5·11) on the positive real axes of ν and η. (ν_0, η_0) is therefore the required saddle point. For the general case $n \neq \frac{1}{2}N$, equations (5·10) can be solved only by numerical methods. It is easy to ascertain that a set of solutions exists which may be employed as the saddle point and that this solution lies on the positive real axes of ν and η. When the saddle point is determined, (5·12) gives us $\log I_1$ immediately.

---

† For unsuitably chosen values of $2X/N$, (5·16) has no positive solution for η. These values of $2X/N$ are impossible to attain in any arrangement of the particles.

The integration of I_2 is similar. For the case $n = \frac{1}{2}N$, the saddle point is the same, i.e. ν_0 and η_0. For the general case $n \neq \frac{1}{2}N$, we look for the saddle point by solving numerically equations (5·10) with the suffix (λ_1) replaced by (λ_2). When the saddle point is determined, $\log I_2$ follows at once. Adding I_1 and I_2 together, we get the required $h(N, n, X)$. At $n = \frac{1}{2}N$, $\log I_1$ and $\log I_2$ are given by the same expression apart from terms of the order unity (for the saddle points are the same), so $\log h$ is given by the common expression of $\log I_1$ and $\log I_2$. When $\log I_1$ and $\log I_2$ are not given by the same expression, as in fact this is the case at $n \neq \frac{1}{2}N$, the smaller one of I_1, I_2 is generally negligible compared with the larger (owing to the fact that $\log I_1$ and $\log I_2$ are of the order N), so $\log h$ or $\log(I_1 + I_2)$ is simply the larger of $\log I_1$ and $\log I_2$.

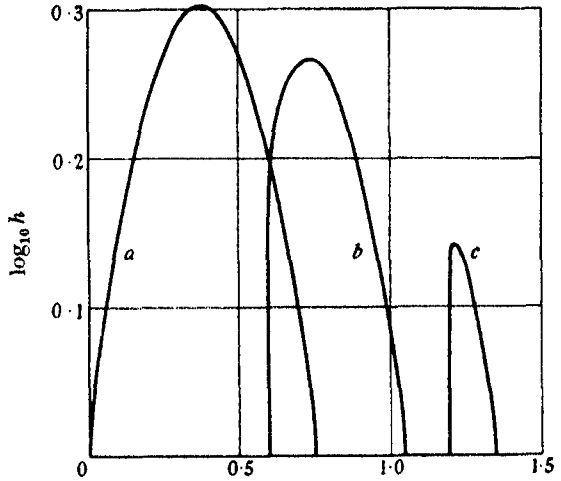


FIGURE 1. $\log_{10} h(N, n, X)$ plotted against X/N. Curve (a) $n/N = 0{\cdot}50$; (b) $n/N = 0{\cdot}70$; (c) $n/N = 0{\cdot}90$.

The results for $N^{-1}\log_{10} h(N, n, X)$ obtained here are plotted below for $n/N = 0{\cdot}50$, $n/N = 0{\cdot}70$, and $n/N = 0{\cdot}90$ (figure 1). Curves with $n/N < 0{\cdot}50$ can be omitted since

$$\log h(N, n, X) = \log h(N, N - n, \tfrac{3}{2}N - 3n + X) \qquad (5{\cdot}17)$$

apart from a possible difference in terms of the order unity.

In the present arrangement of the N points, z is equal to 3. So the formula (1·1) with z put as 3 gives an approximate expression for $\log h(N, n, X)$.

For the sake of comparison, $N^{-1}\log_{10}h(N,n,X)$ as given by (1·1) with z put as 3 is plotted below, for $n/N = 0{\cdot}50, 0{\cdot}70, 0{\cdot}90$ (figure 2). The agreement between corresponding curves in the two figures is bad regarding the edges of the curves, but fairly good regarding their maximums and central parts.

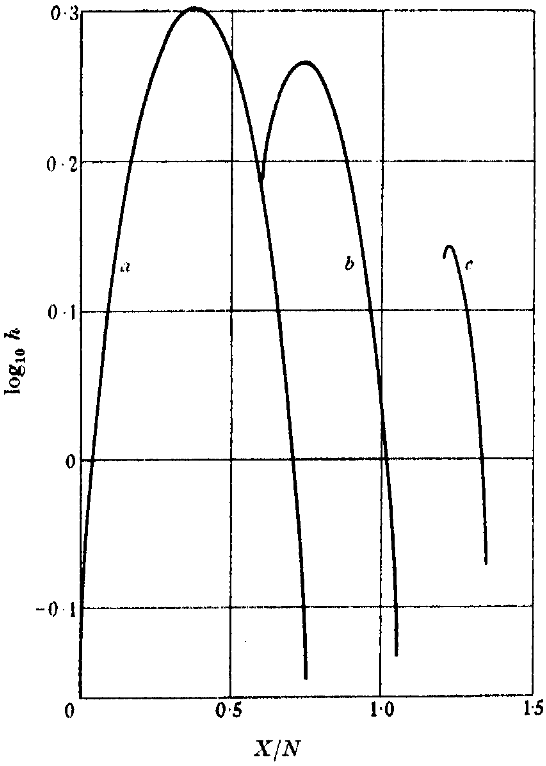


FIGURE 2. Approximate formula for $\log_{10}h(N, n, X)$, i.e. formula (1·1) with $y=3$, plotted against X/N. (a) $n/N=0{\cdot}50$; (b) $n/N=0{\cdot}70$; (c) $n/N=0{\cdot}90$.

It must be mentioned that $h(N,n,X)$ obtained here corresponds to $g_t(N,n,X)$ (or total $g(N,n,X)$) for $z = 3$ in reference 1. In that paper we have described the function $N^{-1}\log g_t(N,n,X)$ for $z = 3$ as a function of n/N and X/N with discontinuities in its partial derivative to X/N at certain values of n/N and X/N. The expectation of such discontinuities is based upon the belief that when X/N is very large, a large part of $g_t(N,n,X)$ ways of arrangements are those in which a part of the N points are crowded with particles and another part empty, and when X/N is very small, a large part of the arrangements are those in which some regularities in the arrangement of the particles appear, resulting in different behaviour of $N^{-1}\log g_t(N,n,X)$ at large and small values of X/N and consequently a possible discontinuity in its derivative to X/N. According to reference 1, (1·1) is the value of $\log g_t(N,n,X)$ only at intermediate values of X/N, and from it we can actually derive a discontinuity of the nature described above when $z > 2$. But here we find no discontinuities for the derivative of $N^{-1}\log h(N, n, X)$ to X/N. This is certainly a disagreement, but not a serious set-back, for the area (or volume) enclosing the N points extends to infinity along one direction only as N approaches infinity in the present case, while in reference 1 it may extend to infinity in several directions.

It must not be thought that there is no possibility of getting such discontinuities in the present calculation. Discontinuities of the nature described above can actually arise if I_1 is greater than I_2 at some values of X/N and smaller at some other, implying that $\log h(N, n, X)$ is given by $\log I_1$ at some X/N and by $\log I_2$ at some other X/N.

6. GENERALIZATIONS AND CONCLUSIONS

It will not be difficult to see now that the method can be extended to study the number of ways of arranging any number of kinds of particles on a number of parallel planes containing in each a finite number of them, given both the numbers of these particles and the different pairs of neighbours formed by them. For we may represent several adjacent planes with their points occupied in a certain manner as a figure and then all configurations are permutations of different kinds of such figures along a straight line. The Φ function in our problem can then be obtained from the $\Psi^{(q)}$'s by performing a substitution of the variables. With Φ obtained, we perform an integration and obtain the number required.

It is not a necessary condition for the method to be applicable that the number of points in each figure or their arrangement be the same. When the number of points or their arrangement in the figures are different, it is possible that some of the ways of permuting the figures may not be compatible with the given arrangement of the N points. So in calculating Φ and $\Psi^{(q)}$, some terms have to be excluded from $\Psi^{(q)}$, which can be generally done by making some of the η's in $\Psi^{(q)}$ zero.

Nor is it necessary that we should restrict ourselves to nearest or next nearest neighbours in specifying the number of ways of arrangements. Given the number of any kind of neighbours, all we have to do in finding the corresponding number of ways of arrangements is to let the figures be so great that there does not exist any neighbour of the given kind formed by particles from two figures separated by a third one lying between them.

Finally, it is not necessary that each particle must occupy one and only one point, as is so understood in the above. This will be illustrated in a paper in course of preparation which investigates the number of ways of arranging a certain number of double particles on two rows of points, each double particle occupying two adjacent points instead of one.

In carrying out the calculation, we meet invariably a determinant and a set of equations such as (4·6), (5·10), etc. The order of the determinant increases rapidly with the number of points in the figures and its evaluation offers the greatest obstacle to the application of the method. When the determinant is evaluated, the problem can be considered as solved, since we can always solve equations such as (5·10) by numerical methods, the rest being simply a substitution into expressions such as (5·12).

Our ultimate purpose is to find $g(N, n, X)$ with the N points lying on a lattice, which may be considered to extend to infinity in all directions. The number of points in each figure will be infinite and we met accordingly a determinant of infinite numbers of rows and columns. If we can only evaluate this determinant, we can obtain our required number. The evaluation, or rather the asymptotic evaluation, of this infinite determinant is therefore the only difficulty left, but we are not optimistic in hoping for its removal. The matter is still in course of investigation.

In conclusion, the writers wish to thank their colleague, Mr S.K. Kao, for much help he has generously given during the preparation of the paper, and members of the department for their kind interest.

REFERENCES

Chang, T. S. 1939*a* *Proc. Roy. Soc.* A, **173**, 48.
Chang, T. S. 1939*b* *Proc. Camb. Phil. Soc.* **35**, 265.
Fowler, R. H. 1936 *Statistical Mechanics*, 2nd ed. Cambridge.
Ising, E. 1925 *Z. Phys.* **31**, 253.

Some practically important stress-systems in solids of revolution

By R. V. Southwell, F.R.S.

(*Received* 19 *November* 1941)

The 'semi-indirect' method of Saint Venant is applied to determine directly (i.e. without explicit reference to strains) distributions of stress in complete solids of revolution (e.g. twisted shafts of non-uniform diameter, axially symmetrical systems such as pressure vessels) and in incomplete tores (e.g. helical springs of small pitch, beams of large curvature, and hooks). The paper is intended as a prelude to approximate but general treatment by relaxation methods; consequently only simple examples are considered, but a comparison is made with earlier solutions less suited to a 'relaxation' approach.

INTRODUCTION AND SUMMARY

This paper records investigations started more than 20 years ago, but not then brought to publication for the reason that their boundary conditions seemed to oppose insuperable difficulties to an orthodox attack. Recent extensions of the Relaxation Method have gone far to remove those difficulties, and lately trial has shown them to be in fact surmountable. Consequently the analysis has now acquired utility, and it is set out here as a background to papers in course of preparation.

Other investigators (notably Michell 1899) have attacked the same problems, and in some instances by treatment similarly aimed at direct determination of the stresses, without explicit reference to strain. Since each of the six stress components is biharmonic in Cartesian co-ordinates x, y, z, and since (in Love's notation)

$$\Theta = X_x + Y_y + Z_z$$

is harmonic in the same co-ordinates, so that

$$\nabla^2\Theta = 0,$$

it was not to be expected that a new investigation would lead to substantially different results. But although the stress functions of this paper have general similarity with those of the earlier solutions, they do exhibit certain differences, and in particular their boundary conditions are better suited to a 'relaxation' approach. For orthodox analysis their advantages are probably not appreciable; but orthodox analysis has not in many cases, so far, led to practically useful results.

Applications are not described for the present, since the purpose of this paper is to provide references for future work. But it may perhaps have some theoretical interest of its own, as a systematic examination of those cases in which, for cylindrical co-ordinates r, θ, z, the general problem of elastic equilibrium can be reduced by reasonable assumptions to one involving only two independent variables.

The cases treated are five in number, viz.

Problem (i, 1) *a*: Torsional stresses in an incomplete tore (§§ 9–10).

Problem (i, 1) *b*: Torsion of a circular shaft having non-uniform diameter (§ 11).

Problem (i, 2) *a*: Symmetrical strain in a solid of revolution (§§ 16–17).

Problem (i, 2) *b*: Flexural stresses in an incomplete tore (§§ 18–19).

Problem (ii): Shearing and flexural stresses in a toroidal 'hook' (§§ 24–27).

To each a formal solution is obtained, illustrated by a simple example,* and compared with earlier solutions. It does not appear that any general treatment of Problem (ii) has been given previously.

I. GENERAL THEORY

The governing equations

1. The stress components in an isotropic elastic body, in equilibrium and free from body forces, satisfy three stress equations of equilibrium, typified by

$$\frac{\partial X_x}{\partial x}+\frac{\partial X_y}{\partial y}+\frac{\partial X_z}{\partial z}=0, \tag{1}$$

and six 'conditions of compatibility',† typified by

$$\nabla^2 X_x+\frac{1}{1+\sigma}\frac{\partial^2\Theta}{\partial x^2}=0 \tag{2}$$

and by

$$\nabla^2 Y_z+\frac{1}{1+\sigma}\frac{\partial^2\Theta}{\partial y\,\partial z}=0, \tag{3}$$

in which σ stands for Poisson's ratio and $\Theta = X_x+Y_y+Z_z$, so that

$$\nabla^2\Theta=0 \tag{4}$$

according to the three equations typified by (2).

---

* With the exception of Problem (i, 1) *a*: cf. § 28.

† These must be satisfied in order that the accompanying strains may be compatible with single-valued displacements.

For the purpose of this paper cylindrical co-ordinates (r, θ, z) are more appropriate than Cartesian (x, y, z). In this system the stress equations (1) are replaced (Love 1927, § 59) by

$$\left.\begin{aligned}
&\frac{\partial \widehat{rr}}{\partial r}+\frac{1}{r}\frac{\partial \widehat{r\theta}}{\partial \theta}+\frac{\partial \widehat{rz}}{\partial z}+\frac{\widehat{rr}-\widehat{\theta\theta}}{r}=0,\\
&\frac{\partial \widehat{r\theta}}{\partial r}+\frac{1}{r}\frac{\partial \widehat{\theta\theta}}{\partial \theta}+\frac{\partial \widehat{\theta z}}{\partial z}+2\frac{\widehat{r\theta}}{r}=0,\\
&\frac{\partial \widehat{rz}}{\partial r}+\frac{1}{r}\frac{\partial \widehat{\theta z}}{\partial \theta}+\frac{\partial \widehat{zz}}{\partial z}+\frac{\widehat{rz}}{r}=0,
\end{aligned}\right\} \qquad (5)$$

and equations of compatibility can be deduced without difficulty from the formulae for transformation of stress components (Love 1927, § 49), the scheme being

	r	θ	z
x	$\cos\theta$	$-\sin\theta$	0
y	$\sin\theta$	$\cos\theta$	0
z	0	0	1

, (6)

so that (for example)

$$\left.\begin{aligned}
\frac{\partial^2}{\partial x^2}&\equiv\frac{1}{2}\left(\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}\right)+\tfrac{1}{2}\cos 2\theta\left(\frac{\partial^2}{\partial r^2}-\frac{1}{r}\frac{\partial}{\partial r}-\frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}\right)\\
&\qquad\qquad -\sin 2\theta\left(\frac{1}{r}\frac{\partial^2}{\partial r\,\partial\theta}-\frac{1}{r^2}\frac{\partial}{\partial\theta}\right),\\
\frac{\partial^2}{\partial x\,\partial y}&\equiv\tfrac{1}{2}\sin 2\theta\left(\frac{\partial^2}{\partial r^2}-\frac{1}{r}\frac{\partial}{\partial r}-\frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}\right)\\
&\qquad\qquad +\cos 2\theta\left(\frac{1}{r}\frac{\partial^2}{\partial r\,\partial\theta}-\frac{1}{r^2}\frac{\partial}{\partial\theta}\right),\\
\text{and}\qquad \nabla^2&\equiv\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}+\frac{\partial^2}{\partial z^2}.
\end{aligned}\right\} \qquad (7)$$

2. From the formulae of transformation we obtain

$$\left.\begin{aligned}
X_x &= \tfrac{1}{2}(\widehat{rr}+\widehat{\theta\theta})+\tfrac{1}{2}(\widehat{rr}-\widehat{\theta\theta})\cos 2\theta-\widehat{r\theta}\sin 2\theta,\\
Y_y &= \tfrac{1}{2}(\widehat{rr}+\widehat{\theta\theta})-\tfrac{1}{2}(\widehat{rr}-\widehat{\theta\theta})\cos 2\theta+\widehat{r\theta}\sin 2\theta,\\
Z_z &= \widehat{zz},\quad Y_z=\widehat{\theta z}\cos\theta+\widehat{zr}\sin\theta,\\
Z_x &= \widehat{zr}\cos\theta-\widehat{\theta z}\sin\theta,\\
X_y &= \tfrac{1}{2}(\widehat{rr}-\widehat{\theta\theta})\sin 2\theta+\widehat{r\theta}\cos 2\theta,
\end{aligned}\right\} \qquad (8)$$

so that

$$\Theta=\widehat{rr}+\widehat{\theta\theta}+\widehat{zz} \qquad (9)$$

satisfies (4) as before. Then from (2) and (3), which must be satisfied identically by (8), we have

$$\left.\begin{aligned}
&\nabla^2(\widehat{rr}+\widehat{\theta\theta})+\left[\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}\right]\left(\frac{\Theta}{1+\sigma}\right)=0,\\
&\left[\nabla^2-\frac{4}{r^2}\right](\widehat{rr}-\widehat{\theta\theta})-\frac{8}{r^2}\frac{\partial}{\partial\theta}\widehat{r\theta}+\left[\frac{\partial^2}{\partial r^2}-\frac{1}{r}\frac{\partial}{\partial r}-\frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}\right]\left(\frac{\Theta}{1+\sigma}\right)=0,\\
&\left[\nabla^2-\frac{4}{r^2}\right]\widehat{r\theta}+\frac{2}{r^2}\frac{\partial}{\partial\theta}(\widehat{rr}-\widehat{\theta\theta})+\left[\frac{1}{r}\frac{\partial^2}{\partial r\,\partial\theta}-\frac{1}{r^2}\frac{\partial}{\partial\theta}\right]\left(\frac{\Theta}{1+\sigma}\right)=0,\\
&\nabla^2\widehat{zz}+\frac{\partial^2}{\partial z^2}\left(\frac{\Theta}{1+\sigma}\right)=0,\\
&\left[\nabla^2-\frac{1}{r^2}\right]\widehat{zr}-\frac{2}{r^2}\frac{\partial}{\partial\theta}\widehat{\theta z}+\frac{\partial^2}{\partial z\,\partial r}\left(\frac{\Theta}{1+\sigma}\right)=0,\\
\text{and}\quad&\frac{2}{r^2}\frac{\partial}{\partial\theta}\widehat{zr}+\left[\nabla^2-\frac{1}{r^2}\right]\widehat{\theta z}+\frac{1}{r}\frac{\partial^2}{\partial\theta\,\partial z}\left(\frac{\Theta}{1+\sigma}\right)=0.
\end{aligned}\right\}\qquad(10)$$

These are the conditions of compatibility expressed in cylindrical coordinates.

The boundary conditions

3. The boundaries which concern us in this paper are surfaces generated by rotation about Oz of curves such as AB, CD, LMN (figure 1). The traction at any point on a boundary may be specified by its components $\widehat{\nu\nu}$, $\widehat{\nu s}$ and $\widehat{\nu\theta}$, θ denoting as before a direction perpendicular to the axial plane, ν the outward-drawn normal to the generating curve, and s being measured along that curve in a direction such that ν, s, θ are related with z, r, θ by the scheme of transformation

	z	r	θ
ν	$\sin(r,\nu)$	$\cos(r,\nu)$	0
s	$-\cos(r,\nu)$	$\sin(r,\nu)$	0
θ	0	0	1

. (11)

Hence we have the relations

$$\left.\begin{aligned}
\widehat{\nu r}&=\widehat{\nu\nu}\cos(r,\nu)+\widehat{\nu s}\sin(r,\nu)=\widehat{rr}\cos(r,\nu)+\widehat{zr}\sin(r,\nu),\\
\widehat{\nu\theta}&=\widehat{r\theta}\cos(r,\nu)+\widehat{\theta z}\sin(r,\nu)\quad\text{(as before)},\\
\widehat{\nu z}&=\widehat{\nu\nu}\sin(r,\nu)-\widehat{\nu s}\cos(r,\nu)=\widehat{zr}\cos(r,\nu)+\widehat{zz}\sin(r,\nu),
\end{aligned}\right\}\qquad(12)$$

as boundary conditions in our problems.

(i) *Stress systems which are independent of θ*

4. Equations (5), (10) and (12) are entirely unrestricted, and we shall not attempt their general solution in this paper. Instead, by restrictive assumptions which reduce the number of the independent variables from three to two, we now derive from them governing equations for several problems which have practical importance.

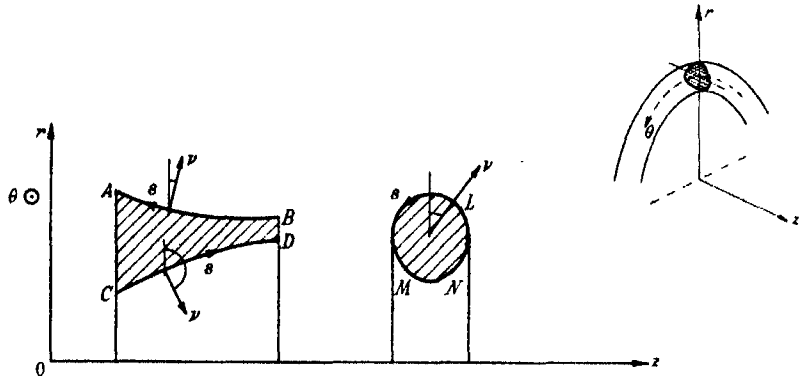


FIGURE 1

Thus assuming that all stress components are independent of θ, we have

$$\left.\begin{aligned} &\frac{\partial\widehat{rr}}{\partial r}+\frac{\partial\widehat{zr}}{\partial z}+\frac{\widehat{rr}-\widehat{\theta\theta}}{r}=0,\\ &\frac{\partial\widehat{r\theta}}{\partial r}+\frac{\partial\widehat{\theta z}}{\partial z}+2\frac{\widehat{r\theta}}{r}=0,\\ \text{and}\qquad &\frac{\partial\widehat{zr}}{\partial r}+\frac{\partial\widehat{zz}}{\partial z}+\frac{\widehat{zr}}{r}=0, \end{aligned}\right\}\tag{13}$$

as simplified forms of (5), and

$$\left.\begin{aligned} &\nabla^2(\widehat{rr}+\widehat{\theta\theta})+\left[\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}\right]\left(\frac{\Theta}{1+\sigma}\right)=0,\\ &\left[\nabla^2-\frac{4}{r^2}\right](\widehat{rr}-\widehat{\theta\theta})+\left[\frac{\partial^2}{\partial r^2}-\frac{1}{r}\frac{\partial}{\partial r}\right]\left(\frac{\Theta}{1+\sigma}\right)=0,\\ &\left[\nabla^2-\frac{4}{r^2}\right]\widehat{r\theta}=0,\\ &\nabla^2\widehat{zz}+\frac{\partial^2}{\partial z^2}\left(\frac{\Theta}{1+\sigma}\right)=0,\\ &\left[\nabla^2-\frac{1}{r^2}\right]\widehat{zr}+\frac{\partial^2}{\partial z\,\partial r}\left(\frac{\Theta}{1+\sigma}\right)=0,\\ &\left[\nabla^2-\frac{1}{r^2}\right]\widehat{\theta z}=0, \end{aligned}\right\}\tag{14}$$

as simplified forms of (10). As before

$$\Theta = \widehat{rr}+\widehat{\theta\theta}+\widehat{zz}, \tag{9 bis}$$

but now in place of the third of (7) we have

$$\nabla^2 \equiv \frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+\frac{\partial^2}{\partial z^2}. \tag{15}$$

The boundary conditions (12) are not altered, but the applied tractions must of course be independent of θ.

(i, 1). *Stress systems entailing equivoluminal strain*

5. We observe that in (13) and (14), as also in (12), $\widehat{r\theta}$ and $\widehat{\theta z}$ are entirely independent of $\widehat{rr}$, $\widehat{\theta\theta}$, $\widehat{zz}$ and $\widehat{zr}$. Consequently further simplification may be effected by making one of two assumptions: either that $\widehat{rr}$, $\widehat{\theta\theta}$, $\widehat{zz}$, $\widehat{zr}$ vanish everywhere, so that the strain is equivoluminal; or that $\widehat{\theta z}$, $\widehat{r\theta}$ vanish everywhere, so that axial sections remain plane.

Making the first assumption, we postulate that

$$\widehat{rr} = \widehat{\theta\theta} = \widehat{zz} = \widehat{zr} = 0, \tag{16}$$

everywhere. The equations (13) reduce to one, viz.

$$\frac{\partial}{\partial r}(r^2 . \widehat{r\theta})+\frac{\partial}{\partial z}(r^2 . \widehat{\theta z}) = 0, \tag{17}$$

and of the six equations (14), since $\Theta = 0$ by (16), only two are not satisfied identically, viz.

$$\left[\nabla^2-\frac{4}{r^2}\right]\widehat{r\theta} = 0, \quad \left[\nabla^2-\frac{1}{r^2}\right]\widehat{\theta z} = 0. \tag{18}$$

The boundary conditions (12) of § 3 reduce to

$$\left.\begin{aligned} &\widehat{\nu r} = 0, \quad \widehat{\nu z} = 0, \\ \text{with} \quad &\widehat{\nu\theta} = \widehat{r\theta}\cos(r,\nu)+\widehat{\theta z}\sin(r,\nu), \quad \text{as before.}\end{aligned}\right\} \tag{19}$$

(i, 2). *Stress systems in which no tangential stress is operative on axial planes.*

6. Making the second assumption of § 5, we simplify the general equations on the assumption that

$$\widehat{\theta z} = \widehat{r\theta} = 0. \tag{20}$$

Then, of the three equations (13), the first and last are unaltered but the second is satisfied identically. Of (14), the third and sixth are satisfied identically, the remainder are unaltered. Of the boundary conditions (12), the second is satisfied identically.

(ii) ***Systems in which the stress components have*** $\sin\theta$ ***or*** $\cos\theta$ ***as a factor.***

7. The simplifying assumptions of §§4–6 have reduced the general problem to problems involving only two independent variables r, z. Another assumption leading to a like result is that *in so far as the stress components depend on* θ

$$\widehat{rr},\ \widehat{\theta\theta},\ \widehat{zz},\ \widehat{zr} \propto \sin\theta, \quad \widehat{r\theta},\ \widehat{\theta z} \propto \cos\theta. \tag{21}$$

On this understanding, and *with omission of the factors in* θ, equations (5) reduce to

$$\left.\begin{aligned}
&\frac{\partial \widehat{rr}}{\partial r} - \frac{\widehat{r\theta}}{r} + \frac{\partial \widehat{rz}}{\partial z} + \frac{\widehat{rr} - \widehat{\theta\theta}}{r} = 0,\\
&\frac{\partial \widehat{r\theta}}{\partial r} + \frac{\widehat{\theta\theta}}{r} + \frac{\partial \widehat{\theta z}}{\partial z} + 2\frac{\widehat{r\theta}}{r} = 0,\\
&\frac{\partial \widehat{rz}}{\partial r} - \frac{\widehat{\theta z}}{r} + \frac{\partial \widehat{zz}}{\partial z} + \frac{\widehat{rz}}{r} = 0,
\end{aligned}\right\} \tag{22}$$

and equations (10) to

$$\left.\begin{aligned}
&\nabla^2(\widehat{rr} + \widehat{\theta\theta}) + \left[\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} - \frac{1}{r^2}\right]\left(\frac{\Theta}{1+\sigma}\right) = 0,\\
&\left[\nabla^2 - \frac{4}{r^2}\right](\widehat{rr} - \widehat{\theta\theta}) + 8\frac{\widehat{r\theta}}{r^2} + \left[\frac{\partial^2}{\partial r^2} - \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\right]\left(\frac{\Theta}{1+\sigma}\right) = 0,\\
&\left[\nabla^2 - \frac{4}{r^2}\right]\widehat{r\theta} + \frac{2}{r^2}(\widehat{rr} - \widehat{\theta\theta}) + \left[\frac{1}{r}\frac{\partial}{\partial r} - \frac{1}{r^2}\right]\left(\frac{\Theta}{1+\sigma}\right) = 0,\\
&\nabla^2\widehat{zz} + \frac{\partial^2}{\partial z^2}\left(\frac{\Theta}{1+\sigma}\right) = 0,\\
&\left[\nabla^2 - \frac{1}{r^2}\right]\widehat{zr} + \frac{2}{r^2}\widehat{\theta z} + \frac{\partial^2}{\partial z\,\partial r}\left(\frac{\Theta}{1+\sigma}\right) = 0,\\
&\frac{2}{r^2}\widehat{zr} + \left[\nabla^2 - \frac{1}{r^2}\right]\widehat{\theta z} + \frac{1}{r}\frac{\partial}{\partial z}\left(\frac{\Theta}{1+\sigma}\right) = 0,
\end{aligned}\right\} \tag{23}$$

Θ now standing for the quantity $(\widehat{rr} + \widehat{\theta\theta} + \widehat{zz})$ *with its factor* $\sin\theta$ *omitted*, and

$$\nabla^2 \text{ now denoting } \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} - \frac{1}{r^2} + \frac{\partial^2}{\partial z^2}. \tag{24}$$

In the boundary conditions (§3) our assumption requires that

$$\widehat{\nu r},\ \widehat{\nu z} \propto \sin\theta, \quad \widehat{\nu\theta} \propto \cos\theta. \tag{25}$$

The forms of the boundary conditions (12) are unaltered if the factors $\sin\theta$, $\cos\theta$ are again omitted.

II. FORMAL SOLUTIONS

8. We proceed to derive formal solutions of the simplified problems of §§ 5–7, and to consider their practical application. Discussion from the standpoint of relaxation methods is reserved for a subsequent paper.

Case (i, 1), § 5.

Here the equations to be solved are (17)–(19) of § 5. Combining (17) and (18), we deduce that

$$r^2 . \widehat{r\theta} = \frac{\partial \phi}{\partial z}, \quad r^2 . \widehat{\theta z} = -\frac{\partial \phi}{\partial r}, \tag{26}$$

where ϕ is some function satisfying the equation

$$\left[\frac{\partial^2}{\partial r^2} - \frac{3}{r}\frac{\partial}{\partial r} + \frac{\partial^2}{\partial z^2}\right]\phi = \text{const.} = A \text{ (say)}. \tag{27}$$

The first and second of (19) require that

$$\widehat{\nu r} = \widehat{\nu z} = 0, \tag{28}$$

so that the boundary traction (*if non-zero*) *is everywhere tangential and directed with* θ. It is given by the third of (19), i.e. by

$$\begin{aligned} r^2 . \widehat{\nu\theta} &= \cos(r, \nu)\frac{\partial \phi}{\partial z} - \sin(r, \nu)\frac{\partial \phi}{\partial r} \\ &= -\frac{\partial \phi}{\partial s} \text{ (cf. figure 1)}. \end{aligned} \tag{29}$$

Accordingly on an unloaded boundary ($\widehat{\nu\theta}$ being zero) we have

$$\phi = \text{const.} = C \text{ (say)}. \tag{30}$$

Torsional stresses in an incomplete tore (*Problem* (i, 1) *a*)

9. A case of some practical importance is obtained when the boundary is unloaded and consists of a single closed curve (*LMN*, figure 1). Here the constant in (30) is immaterial, since it implies—according to (27)—a constant addition to ϕ, which will not affect the stress components. Accordingly we may replace (30) by

$$\phi = 0, \quad \text{on the boundary}. \tag{31}$$

This, with (27), defines our problem.

10. The stresses on a cross-section have a resultant

$$\begin{aligned} \mathbf{R} = \iint \widehat{r\theta}\, dr\, dz = \iint \frac{\partial}{\partial z}\left(\frac{\phi}{r^2}\right) dr\, dz = \oint \sin(r, \nu)\frac{\phi}{r^2} ds \\ = 0, \quad \text{by (31)}, \end{aligned}$$

in the direction of r, and a resultant

$$\mathbf{Z} = \iint \widehat{\theta z}\, dr\, dz = -\iint \frac{1}{r^2}\frac{\partial\phi}{\partial r}\, dr\, dz$$

in the direction of z. In virtue of (27) we may replace this second expression by

$$\begin{aligned}\mathbf{Z} &= \frac{1}{2}\iint\left\{\frac{A}{r} - \frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial\phi}{\partial r}\right) - \frac{\partial}{\partial z}\left(\frac{1}{r}\frac{\partial\phi}{\partial z}\right)\right\} dr\, dz\\ &= \frac{1}{2}\oint\left\{\cos(r,\nu)\left(A\log r - \frac{1}{r}\frac{\partial\phi}{\partial r}\right) - \sin(r,\nu)\frac{1}{r}\frac{\partial\phi}{\partial z}\right\} ds\\ &= \frac{1}{2}\oint\left\{\cos(r,\nu)\, A\log r - \frac{1}{r}\frac{\partial\phi}{\partial\nu}\right\} ds\\ &= \tfrac{1}{2}A\oint\left\{\cos(r,\nu)\log r - \frac{1}{r}\frac{\partial\phi}{\partial\nu}\right\} ds \qquad (32)\end{aligned}$$

if we now write $A\phi$ for the function hitherto denoted by ϕ, so that ϕ still satisfies (31) but is now governed by

$$\left[\frac{\partial^2}{\partial r^2} - \frac{3}{r}\frac{\partial}{\partial r} + \frac{\partial^2}{\partial z^2}\right]\phi = 1. \qquad (27a)$$

On that understanding the moment of the shearing stresses about the origin ($r = 0$, $z = 0$) is measured by

$$\begin{aligned}\iint (z\,.\,\widehat{r\theta} - r\,.\,\widehat{\theta z})\, dr\, dz &= A\iint\left\{z\frac{\partial}{\partial z}\left(\frac{\phi}{r^2}\right) + \frac{1}{r}\frac{\partial\phi}{\partial r}\right\} dr\, dz\\ &= A\iint\left\{\frac{\partial}{\partial r}\left(\frac{\phi}{r}\right) + \frac{\partial}{\partial z}\left(\frac{z\phi}{r^2}\right)\right\} dr\, dz\\ &= A\oint\{r\cos(r,\nu) + z\sin(r,\nu)\}\frac{\phi}{r^2}\, ds\\ &= 0, \quad \text{by (31)}.\end{aligned}$$

Consequently the only action on a cross-section is a force $\mathbf{Z}$ given by (32), with line of action Oz. Knowing $\mathbf{Z}$ we can use (32) to determine A.

Thus the solution of case (i, 1) *a* gives a stress system which would result from shearing actions, suitably distributed over the terminal sections, of which the resultants are directed along the axis of the tore. Actions of this kind are exerted when a helical spring of small pitch withstands tension or compression, and in practice it is customary to regard the coiled wire as subjected to pure torsion and to compute the stresses by the torsion theory of Saint Venant. Our solution is a closer approximation, making allowance for the wire's initial curvature by the term $-\frac{3}{r}\frac{\partial\phi}{\partial r}$ in (27). It will be for all practical purposes exact when the pitch of the helix is small.

Torsion of circular shaft with non-uniform diameter (*Problem* (i, 1) *b*)

11. In order that A in (27) may be non-zero, the elastic body must be an *incomplete* solid of revolution. This can be demonstrated as follows: If u, v, w are component displacements in the directions of r, θ, z, the stress-strain relations require that

$$\left.\begin{aligned} \frac{1}{r^3}\frac{\partial\phi}{\partial z} &= \frac{\widehat{r\theta}}{r} = \frac{\mu}{r}e_{r\theta} = \mu\left\{\frac{\partial}{\partial r}\left(\frac{v}{r}\right)+\frac{1}{r^2}\frac{\partial u}{\partial\theta}\right\}, \\ -\frac{1}{r^3}\frac{\partial\phi}{\partial r} &= \frac{\widehat{\theta z}}{r} = \frac{\mu}{r}e_{\theta z} = \mu\left\{\frac{1}{r^2}\frac{\partial w}{\partial\theta}+\frac{\partial}{\partial z}\left(\frac{v}{r}\right)\right\}, \end{aligned}\right\} \qquad \text{(i)}$$

μ denoting the modulus of rigidity. Hence, eliminating v/r, we have

$$\mu\frac{\partial}{\partial\theta}\left\{\frac{1}{r^2}\frac{\partial u}{\partial z}-\frac{\partial}{\partial r}\left(\frac{w}{r^2}\right)\right\} = \frac{\partial}{\partial r}\left(\frac{1}{r^3}\frac{\partial\phi}{\partial r}\right)+\frac{\partial}{\partial z}\left(\frac{1}{r^3}\frac{\partial\phi}{\partial z}\right) = \frac{A}{r^3} \qquad \text{(ii)}$$

according to (27); and integrating (ii) with respect to θ we obtain a cyclic expression for the quantity in twisted brackets unless A is zero, so that (27) and (27a) are replaced by

$$\left[\frac{\partial^2}{\partial r^2}-\frac{3}{r}\frac{\partial}{\partial r}+\frac{\partial^2}{\partial z^2}\right]\phi = 0. \qquad (33)$$

Combined with (31) this equation would require the stresses as given by (26) to vanish everywhere, since it implies that

$$\begin{aligned} 0 &= \iint\phi\left[\frac{\partial}{\partial r}\left(\frac{1}{r^3}\frac{\partial\phi}{\partial r}\right)+\frac{\partial}{\partial z}\left(\frac{1}{r^3}\frac{\partial\phi}{\partial z}\right)\right]dr\,dz \\ &= \oint\frac{\phi}{r^3}\frac{\partial\phi}{\partial\nu}ds-\iint\frac{1}{r^3}\left\{\left(\frac{\partial\phi}{\partial r}\right)^2+\left(\frac{\partial\phi}{\partial z}\right)^2\right\}dr\,dz, \end{aligned} \qquad (34)$$

and the contour integral vanishes when (31) is satisfied. Consequently no solution for a *complete* tore is derivable from (33).

When, on the other hand, $\widehat{\nu\theta}$ is non-zero over some part at least of the boundary, the same is true of ϕ as determined by (29), and then non-zero stresses are entailed in the solution to (33). An example of practical importance is the torsion of a circular shaft (either solid or tubular) of which the diameter varies along its length.

Case (i, 2), § 6.

12. Here the equations to be solved are (cf. § 6) the first and third of (13) and the first, second, fourth and fifth of (14), Θ and ∇^2 having the significance given in (9) and (15) respectively. The boundary conditions to be satisfied are the first and third of (12), viz.

$$\left.\begin{aligned} \widehat{\nu r} &= \widehat{rr}\cos(r,\nu)+\widehat{zr}\sin(r,\nu), \\ \widehat{\nu z} &= \widehat{zr}\cos(r,\nu)+\widehat{zz}\sin(r,\nu), \end{aligned}\right\} \qquad (35)$$

$\widehat{\nu r}$ and $\widehat{\nu z}$ being specified.

13. Writing
$$\left.\begin{aligned}\widehat{rr}+\widehat{\theta\theta} &= \frac{1}{r}\frac{\partial\chi}{\partial r},\\ \widehat{rr}-\widehat{\theta\theta} &= r\frac{\partial\psi_1}{\partial r},\end{aligned}\right\} \tag{36}$$

without restriction (as yet) on χ and ψ_1, we have
$$\widehat{zz} = \Theta-(\widehat{rr}+\widehat{\theta\theta}) = \Theta-\frac{1}{r}\frac{\partial\chi}{\partial r}. \tag{37}$$

Then the first and second of (14) become
$$\frac{1}{r}\frac{\partial}{\partial r}\left\{\vartheta^2\chi+r\frac{\partial}{\partial r}\left(\frac{\Theta}{1+\sigma}\right)\right\} = 0,$$
$$r\frac{\partial}{\partial r}\frac{1}{r^2}\left\{\vartheta^2(r^2\psi_1)+r\frac{\partial}{\partial r}\left(\frac{\Theta}{1+\sigma}\right)\right\} = 0,$$

ϑ^2 standing for the operator $\left[\frac{\partial^2}{\partial r^2}-\frac{1}{r}\frac{\partial}{\partial r}+\frac{\partial^2}{\partial z^2}\right]$, (38)

and it follows that we may write
$$-r\frac{\partial}{\partial r}\left(\frac{\Theta}{1+\sigma}\right) = \vartheta^2\{\chi-F_1(z)\},$$

and replace the second of (36) by (39)
$$\widehat{rr}-\widehat{\theta\theta} = r\frac{\partial}{\partial r}\left(\frac{\chi+\phi_1}{r^2}\right)+2\frac{F_1(z)}{r^2},$$

ϕ_1 being any solution of the equation $\vartheta^2\phi = 0$, and $F_1(z)$ any function of z.

The first and second of (14) have been satisfied, and the fourth will be satisfied when $\nabla^2\Theta$ has been made to vanish. The fifth can be replaced by
$$\frac{1}{r}\vartheta^2\left\{r\,.\,\widehat{zr}-\frac{\partial\chi}{\partial z}+F_1'(z)\right\} = 0$$

when we substitute from (39), showing that we may write
$$r\,.\,\widehat{zr} = \frac{\partial}{\partial z}\{\chi-F_1(z)\}+\phi_2, \tag{40}$$

where ϕ_2 stands for any solution of the equation $\vartheta^2\phi = 0$.

14. We turn now to the equations (13) of equilibrium. The third is equivalent to
$$\frac{\partial}{\partial r}(r\,.\,\widehat{rz})+\frac{\partial}{\partial z}(r\,.\,\widehat{zz}) = 0,$$

reducing, when we substitute from (37) and (40), to
$$\frac{\partial\phi_2}{\partial r}+r\,.\,\frac{\partial\Theta}{\partial z} = 0.$$

Accordingly we may write (*ψ being an unrestricted function distinct from ψ_1*)

$$-\phi_2 = \frac{\partial\psi}{\partial z}, \quad \Theta = \frac{1}{r}\frac{\partial\psi}{\partial r}, \tag{41}$$

and then, since $\vartheta^2\phi_2 = 0$ and $\nabla^2\Theta = 0$, we have

$$\frac{\partial}{\partial z}, \frac{\partial}{\partial r}(\vartheta^2\psi) = 0,$$

whence $$\vartheta^2\psi = \text{const.} = 2A \text{ (say)}. \tag{42}$$

Equation (40) can now be replaced by

$$r.\widehat{zr} = \frac{\partial}{\partial z}\{\chi - F_1(z) - \psi\},$$

and the first of (39) by

$$\left.\begin{aligned}(1+\sigma)\vartheta^2\{\chi - F_1(z)\} &= -r\frac{\partial}{\partial r}\frac{1}{r}\frac{\partial\psi}{\partial r} = \frac{\partial^2\psi}{\partial z^2} - \vartheta^2\psi \\ &= \frac{\partial^2\psi}{\partial z^2} - 2A, \quad \text{by (42).}\end{aligned}\right\} \tag{43}$$

This means that we can write

$$\left.\begin{aligned}\chi - F_1(z) &= \frac{\partial^2\lambda}{\partial z^2}, \\ \psi &= Az^2 + (1+\sigma)\vartheta^2\lambda, \\ r.\widehat{zr} &= \frac{\partial}{\partial z}\left[\frac{\partial^2\lambda}{\partial z^2} - Az^2 - (1+\sigma)\vartheta^2\lambda\right],\end{aligned}\right\} \tag{44}$$

where, according to (42), $$\vartheta^4\lambda = 0. \tag{45}$$

15. Modifying our expressions in the light of these results, we have from the first of (36)

$$\widehat{rr} + \widehat{\theta\theta} = \frac{1}{r}\frac{\partial^3\lambda}{\partial r\,\partial z^2},$$

from the second of (39)

$$\widehat{rr} - \widehat{\theta\theta} = r\frac{\partial}{\partial r}\frac{1}{r^2}\left(\frac{\partial^2\lambda}{\partial z^2} + \phi_1\right),$$

from (37) and (41) $$r.\widehat{zz} = \frac{\partial}{\partial r}\left\{(1+\sigma)\vartheta^2\lambda - \frac{\partial^2\lambda}{\partial z^2}\right\},$$

and from the last of (44)

$$r.\widehat{zr} = -\frac{\partial}{\partial z}\left\{(1+\sigma)\vartheta^2\lambda + Az^2 - \frac{\partial^2\lambda}{\partial z^2}\right\}. \tag{46}$$

All of equations (14) are satisfied by these expressions, also the last of (13); but we have still to satisfy the first of (13).

Adding the first and second of (46), we have

$$\widehat{rr} = \frac{\partial}{\partial r}\frac{1}{r}\frac{\partial^2\lambda}{\partial z^2} + \tfrac{1}{2}r\frac{\partial}{\partial r}\left(\frac{\phi_1}{r^2}\right),$$

whence

$$\frac{\partial\widehat{rr}}{\partial r} = \frac{1}{r}\left[\frac{\partial^2}{\partial r^2} - \frac{2}{r}\frac{\partial}{\partial r} + \frac{2}{r^2}\right]\frac{\partial^2\lambda}{\partial z^2} + \frac{1}{2r}\left[\frac{\partial^2}{\partial r^2} - \frac{3}{r}\frac{\partial}{\partial r} + \frac{4}{r^2}\right]\phi_1;$$

from the second of (46) again we have

$$\frac{\widehat{rr} - \widehat{\theta\theta}}{r} = \left[\frac{1}{r^2}\frac{\partial}{\partial r} - \frac{2}{r^3}\right]\left(\frac{\partial^2\lambda}{\partial z^2} + \phi_1\right), \tag{47}$$

and from the last of (46)

$$\frac{\partial\widehat{rz}}{\partial z} = -\frac{1}{r}\frac{\partial^2}{\partial z^2}\left\{(1+\sigma)\vartheta^2\lambda + Az^2 - \frac{\partial^2\lambda}{\partial z^2}\right\}.$$

Adding, we find that the expression on the left of the first of (13) is equivalent to

$$\frac{1}{r}\left\{\frac{1}{2}\left[\frac{\partial^2}{\partial r^2} - \frac{1}{r}\frac{\partial}{\partial r}\right]\phi_1 - 2A - \sigma\vartheta^2\frac{\partial^2\lambda}{\partial z^2}\right\},$$

which accordingly must vanish. Hence, using (45) and the fact that $\vartheta^2\phi_1 = 0$, we deduce that

$$\tfrac{1}{2}\phi_1 + \sigma\vartheta^2\lambda = (A\log r + B)r^2 + C + z(Fr^2 + G) - Az^2, \tag{48}$$

B, C, F, G being constants. Consequently we can replace the second of (46) by

$$\widehat{rr} - \widehat{\theta\theta} = r\frac{\partial}{\partial r}\frac{1}{r^2}\left(\frac{\partial^2\lambda}{\partial z^2} - 2\sigma\vartheta^2\lambda\right) + 2A - \frac{4}{r^2}(C + Gz - Az^2), \tag{49}$$

and combining this equation with the first of (46) we have

$$\left.\begin{aligned}\widehat{rr} &= \frac{\partial}{\partial r}\frac{1}{r}\frac{\partial^2\lambda}{\partial z^2} - \sigma r\frac{\partial}{\partial r}\frac{1}{r^2}\vartheta^2\lambda + A - \frac{2}{r^2}(C + Gz - Az^2),\\ \widehat{\theta\theta} &= \frac{1}{r^2}\frac{\partial^2\lambda}{\partial z^2} + \sigma r\frac{\partial}{\partial r}\frac{1}{r^2}\vartheta^2\lambda - A + \frac{2}{r^2}(C + Gz - Az^2),\end{aligned}\right\} \tag{50}$$

with

$$\left.\begin{aligned}\widehat{zz} &= \frac{1}{r}\frac{\partial}{\partial r}\left\{-\frac{\partial^2\lambda}{\partial z^2} + (1+\sigma)\vartheta^2\lambda + Az^2\right\},\\ \widehat{zr} &= -\frac{1}{r}\frac{\partial}{\partial z}\left\{-\frac{\partial^2\lambda}{\partial z^2} + (1+\sigma)\vartheta^2\lambda + Az^2\right\}.\end{aligned}\right\} \tag*{(46) bis}$$

In all of these expressions ϑ^2 carries the significance given in (38).

Symmetrical strain in a complete solid of revolution (*Problem* (i, 2) *a*)

16. Here as in § 11 we must exclude the possibility of cyclic displacements. In a *complete* solid of revolution (by symmetry) there can be no displacement perpendicular to axial planes, consequently the extension corresponding with $\widehat{\theta\theta}$ is u/r simply, and we have

$$(1+\sigma)\widehat{rr}-\sigma\Theta = E\frac{\partial u}{\partial r} = \frac{\partial}{\partial r}[r\{(1+\sigma)\widehat{\theta\theta}-\sigma\Theta\}],$$

i.e.
$$(\widehat{rr}-\widehat{\theta\theta}) = r\frac{\partial}{\partial r}\left\{\widehat{\theta\theta}-\frac{\sigma}{1+\sigma}\Theta\right\},$$

or
$$2(\widehat{rr}-\widehat{\theta\theta}) = r\frac{\partial}{\partial r}\left\{\widehat{rr}+\widehat{\theta\theta}-(\widehat{rr}-\widehat{\theta\theta})-\frac{2\sigma}{1+\sigma}\Theta\right\}. \tag{51}$$

Substituting from (46) and (49) in this equation, we deduce that $A = 0$ in (46), (49) and (50).

Moreover the terms in C and G may be suppressed (when this is desirable) on the ground that they can be incorporated as additions to λ. For an addition

$$\lambda' = (1+\sigma)(\tfrac{1}{3}B'z^3+C'z^2)-\sigma(B'z+C')r^2\log r$$

gives
$$r\frac{\partial}{\partial r}\frac{1}{r}\frac{\partial\lambda'}{\partial r} = -2\sigma(B'z+C'),$$

$$\frac{\partial^2\lambda'}{\partial z^2} = 2(1+\sigma)(B'z+C'),$$

therefore
$$\vartheta^2\lambda' = 2(B'z+C') = \frac{1}{1+\sigma}\frac{\partial^2\lambda'}{\partial z^2},$$

and so entails (on the whole) no addition to $\widehat{zr}$, $\widehat{zz}$, $(\widehat{rr}+\widehat{\theta\theta})$. To $(\widehat{rr}-\widehat{\theta\theta})$ it makes, according to (49), an addition

$$r\frac{\partial}{\partial r}\frac{1}{r^2}\left(\frac{\partial^2\lambda'}{\partial z^2}-2\sigma\vartheta^2\lambda'\right) = -\frac{4}{r^2}(B'z+C')(1-\sigma),$$

which (with suitable values for B' and C') will neutralize the terms in C and G.

17. Utilizing these results, we have in Problem (i, 2)a, in place of the expressions which end § 15,

$$\left.\begin{aligned}
\widehat{rr}+\widehat{\theta\theta} &= \frac{1}{r}\frac{\partial^3\lambda}{\partial r\partial z^2},\\
\widehat{rr}-\widehat{\theta\theta} &= r\frac{\partial}{\partial r}\frac{1}{r^2}\left(\frac{\partial^2\lambda}{\partial z^2}-2\sigma\vartheta^2\lambda\right),\\
\widehat{zz} &= -\frac{1}{r}\frac{\partial}{\partial r}\left\{\frac{\partial^2\lambda}{\partial z^2}-(1+\sigma)\vartheta^2\lambda\right\},\\
\widehat{zr} &= \frac{1}{r}\frac{\partial}{\partial z}\left\{\frac{\partial^2\lambda}{\partial z^2}-(1+\sigma)\vartheta^2\lambda\right\},\\
\Theta &= \frac{1+\sigma}{r}\frac{\partial}{\partial r}\vartheta^2\lambda,
\end{aligned}\right\} \tag{52}$$

λ denoting a function which satisfies the equation

$$\vartheta^4\lambda = 0, \tag{45 bis}$$

but otherwise is (as yet) unrestricted. ϑ^2 as before carries the significance stated in (38).

The boundary conditions (35) can now be written as

$$\left.\begin{aligned} \widehat{\nu r} &= \frac{\partial}{\partial \nu}\left(\frac{1}{r}\frac{\partial^2\lambda}{\partial z^2}\right) - \sigma r\frac{\partial}{\partial \nu}\left(\frac{1}{r^2}\vartheta^2\lambda\right) - \frac{\sin(r,\nu)}{r}\frac{\partial}{\partial z}\vartheta^2\lambda, \\ -\widehat{\nu z} &= \frac{1}{r}\frac{\partial}{\partial s}\left\{\frac{\partial^2\lambda}{\partial z^2} - (1+\sigma)\vartheta^2\lambda\right\}, \end{aligned}\right\} \tag{53}$$

ν and s having the senses indicated in figure 1.

The second of (53) imposes values at every point of the boundary on

$$\frac{\partial^2\lambda}{\partial z^2} - (1+\sigma)\vartheta^2\lambda = \psi \text{ (say)}, \tag{54}$$

ψ having a significance different from what it had in § 14. The constant of integration is immaterial, and ψ is acyclic because

$$\begin{aligned} \oint r\,.\,\widehat{\nu z}\,ds &= \oint r\{\widehat{zr}\cos(r,\nu) + \widehat{zz}\sin(r,\nu)\}\,ds \\ &= \iint\left\{\frac{\partial}{\partial r}(r\,.\,\widehat{zr}) + \frac{\partial}{\partial z}(r\,.\,\widehat{zz})\right\}dr\,dz, \quad \text{for the whole cross-section,} \\ &= 0 \quad \text{by the third of (13).} \end{aligned}$$

The first of (53) imposes boundary values on

$$\widehat{\nu r} = \frac{\partial}{\partial \nu}\left(\frac{\psi}{r}\right) + \cos(r,\nu)\left[\frac{\partial}{\partial r} + \frac{\sigma}{r}\right]\frac{1}{r}\vartheta^2\lambda. \tag{55}$$

Flexural stresses in an incomplete tore (*Problem* (i, 2) *b*)

18. When the solid of revolution is *not* complete (so that cyclic displacements are not excluded) A is no longer required to vanish in (46), (49) and (50), but the argument of § 16 still justifies suppression of the terms involving C and G.

If in addition we replace λ by μ, where

$$\sigma\mu = \sigma\lambda + \frac{1}{12}Az^4, \tag{i}$$

then $\widehat{zz}$, $\widehat{zr}$ will be left unaltered provided that at the same time we suppress the term Az^2 in the last of (46), since

$$-\frac{\partial^2\mu}{\partial z^2} + (1+\sigma)\vartheta^2\mu = -\frac{\partial^2\lambda}{\partial z^2} + (1+\sigma)\vartheta^2\lambda + Az^2.$$

We shall not thereby affect $(\widehat{rr}+\widehat{\theta\theta})$, which according to (50) is given by

$$\frac{1}{r}\frac{\partial^3\lambda}{\partial r\,\partial z^2};$$

and since
$$\widehat{rr}-\widehat{\theta\theta} = r\frac{\partial}{\partial r}\frac{1}{r^2}\left(\frac{\partial^2\lambda}{\partial z^2}-2\sigma\vartheta^2\lambda\right)+2A\left(1+2\frac{z^2}{r^2}\right), \tag{ii}$$

according to (50) again, we shall leave *all* stresses unaltered if, when we replace λ by μ, at the same time we alter the final term of (ii) to

$$2A\left(1+\frac{1}{\sigma}\frac{z^2}{r^2}\right).$$

Now replacing A by σB in our collected results, *when cyclic displacements are not excluded* we can express the general solution by

$$\left.\begin{aligned}
\widehat{rr}+\widehat{\theta\theta} &= \frac{1}{r}\frac{\partial^3\mu}{\partial r\,\partial z^2},\\
\widehat{rr}-\widehat{\theta\theta} &= r\frac{\partial}{\partial r}\frac{1}{r^2}\left(\frac{\partial^2\mu}{\partial z^2}-2\sigma\vartheta^2\mu\right)+2B\left(\sigma+\frac{z^2}{r^2}\right),\\
\widehat{zz} &= \frac{1}{r}\frac{\partial}{\partial r}\left\{-\frac{\partial^2\mu}{\partial z^2}+(1+\sigma)\vartheta^2\mu\right\},\\
\widehat{zr} &= -\frac{1}{r}\frac{\partial}{\partial z}\left\{-\frac{\partial^2\mu}{\partial z^2}+(1+\sigma)\vartheta^2\mu\right\},\\
\Theta &= \frac{1+\sigma}{r}\frac{\partial}{\partial r}\vartheta^2\mu,
\end{aligned}\right\} \tag{56}$$

μ denoting a function of r and z which satisfies the equation

$$\vartheta^4\mu = 2B, \tag{57}$$

according to (45), when ϑ^2 has the significance stated in (38).

19. A practical application of this solution is presented by a curved bar (incomplete tore) which is strained by stresses applied to its terminal sections only. Here $\widehat{\nu r}$, $\widehat{\nu\theta}$, $\widehat{\nu z}$ vanish severally in the boundary conditions (12), of which the second (cf. § 6) is satisfied already: consequently the boundary conditions in this problem are

$$\left.\begin{aligned}
0 &= \widehat{rr}\cos(r,\nu)+\widehat{zr}\sin(r,\nu)\\
&= \frac{\partial}{\partial\nu}\left(\frac{1}{r}\frac{\partial^2\mu}{\partial z^2}\right)-\sigma r\frac{\partial}{\partial\nu}\left(\frac{1}{r^2}\vartheta^2\mu\right)-\frac{\sin(r,\nu)}{r}\frac{\partial}{\partial z}\vartheta^2\mu+B\cos(r,\nu)\left(\sigma+\frac{z^2}{r^2}\right),\\
&\text{and}\\
0 &= \widehat{zr}\cos(r,\nu)+\widehat{zz}\sin(r,\nu) = -\frac{1}{r}\frac{\partial}{\partial s}\left\{\frac{\partial^2\mu}{\partial z^2}-(1+\sigma)\vartheta^2\mu\right\}.
\end{aligned}\right\} \tag{58}$$

The nature of the postulated action can be seen immediately. For the only stress on a cross-section is $\widehat{\theta\theta}$, and according to the first of (13)

$$\begin{aligned}\iint \widehat{\theta\theta}\,dr\,dz &= \iint \left\{\frac{\partial}{\partial r}(r\,.\,\widehat{rr}) + \frac{\partial}{\partial z}(r\,.\,\widehat{zr})\right\} dr\,dz \\ &= \oint r\{\widehat{rr}\cos(r,\nu) + \widehat{zr}\sin(r,\nu)\}\,ds \\ &= 0 \quad \text{by the first of (58).}\end{aligned}$$

Also

$$\begin{aligned}\iint z\,.\,\widehat{\theta\theta}\,dr\,dz &= \iint \left\{\frac{\partial}{\partial r}(zr\,.\,\widehat{rr} - r^2\,.\,\widehat{zr}) + \frac{\partial}{\partial z}(zr\,.\,\widehat{zr} - r^2\,.\,\widehat{zz})\right\} dr\,dz, \\ &= \oint [zr\{\widehat{rr}\cos(r,\nu) + \widehat{zr}\sin(r,\nu)\} \\ &\qquad\qquad - r^2\{\widehat{zr}\cos(r,\nu) + \widehat{zz}\sin(r,\nu)\}]\,ds \\ &= 0, \quad \text{by the first and second of (58).}\end{aligned}$$

Therefore the action on every section is a couple with axis directed along Oz. Its magnitude is

$$\iint r\,.\,\widehat{\theta\theta}\,dr\,dz = \iint \left\{\sigma r^2 \frac{\partial}{\partial r}\left(\frac{1}{r^2}\vartheta^2\mu\right) + \frac{1}{r}\frac{\partial^2\mu}{\partial z^2} - B\left(\sigma r + \frac{z^2}{r}\right)\right\} dr\,dz$$

according to (56).

Case (ii), § 7.

20. Here no stress component vanishes everywhere, so greater difficulty is presented. The equations to be solved are (22) and (23), with boundary conditions as described in § 7. We shall find it convenient to write

$$\left.\begin{aligned} &\vartheta_1^2 \text{ for the operator } \frac{\partial^2}{\partial r^2} - \frac{1}{r}\frac{\partial}{\partial r} + \frac{\partial^2}{\partial z^2}, \\ &\vartheta_3^2 \text{ for the operator } \frac{\partial^2}{\partial r^2} - \frac{3}{r}\frac{\partial}{\partial r} + \frac{\partial^2}{\partial z^2}. \\ &\text{Then, } \nabla^2 \text{ being defined by (24), we have} \\ &\vartheta_1^2 \frac{1}{r}\frac{\partial}{\partial r} \equiv \frac{1}{r}\frac{\partial}{\partial r}\vartheta_3^2, \quad \nabla^2\frac{1}{r^2}\frac{\partial}{\partial r} \equiv \frac{1}{r^2}\frac{\partial}{\partial r}\vartheta_3^2. \end{aligned}\right\} \tag{59}$$

By addition of the first and second of (22), multiplied by r, we obtain

$$\frac{\partial}{\partial r}\{r(\widehat{rr} + \widehat{r\theta})\} + \frac{\partial}{\partial z}\{r(\widehat{\theta z} + \widehat{zr})\} = 0,$$

which shows that we may write

$$\left.\begin{aligned} -r(\widehat{rr} + \widehat{r\theta}) &= \frac{\partial}{\partial z}\left(\frac{1}{r}\frac{\partial^2\psi}{\partial r\,\partial z}\right), \\ r(\widehat{\theta z} + \widehat{zr}) &= \frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial^2\psi}{\partial r\,\partial z}\right), \end{aligned}\right\} \tag{60}$$

ψ denoting a function of z and r as yet undetermined. Also the last of (22), multiplied by r^2, may be written in the form

$$\frac{\partial}{\partial z}(r^2 . \widehat{zz}) + \frac{\partial}{\partial r}(r^2 . \widehat{zr}) = r(\widehat{\theta z} + \widehat{zr}),$$

and it follows from (60) that we may write

$$\left.\begin{aligned} r^2 . \widehat{zz} &= \frac{\partial \phi}{\partial r}, \\ r^2 . \widehat{zr} &= \frac{\partial}{\partial z}\left(\frac{1}{r}\frac{\partial \psi}{\partial r} - \phi\right), \\ r^2 . \widehat{\theta z} &= \frac{\partial}{\partial z}\left(\left[\frac{\partial^2}{\partial r^2} - \frac{2}{r}\frac{\partial}{\partial r}\right]\psi + \phi\right), \end{aligned}\right\} \tag{61}$$

ϕ denoting another function of z and r as yet undetermined.

21. Eliminating Θ between the second and third of (23), we obtain

$$\left[\frac{\partial^2}{\partial r^2} - \frac{1}{r}\frac{\partial}{\partial r} - \frac{3}{r^2} + \frac{\partial^2}{\partial z^2}\right]\left\{\widehat{rr} - \widehat{\theta\theta} - \frac{\partial}{\partial r}(r . \widehat{r\theta})\right\} = 0,$$

in which the quantity contained within twisted brackets

$$= -r\left\{\frac{\partial}{\partial r}(\widehat{rr} + \widehat{r\theta}) + \frac{\partial \widehat{zr}}{\partial z}\right\} \quad \text{in virtue of the first of (22),}$$

$$= \frac{1}{r}\frac{\partial^2}{\partial z^2}\left\{\left[\frac{\partial^2}{\partial r^2} - \frac{3}{r}\frac{\partial}{\partial r}\right]\psi + \phi\right\}, \quad \text{according to (60) and (61).}$$

Consequently $$\frac{\partial^2}{\partial z^2}\left\{\vartheta_3^2 \phi + \left[\frac{\partial^2}{\partial r^2} - \frac{3}{r}\frac{\partial}{\partial r}\right]\vartheta_3^2 \psi\right\} = 0. \tag{62}$$

Moreover, the last of (23) can be written as

$$\frac{\partial}{\partial z}\left\{\frac{\Theta}{1+\sigma} + \frac{1}{r}\vartheta_3^2 \phi + \frac{1}{r}\left[\frac{\partial^2}{\partial r^2} - \frac{2}{r}\frac{\partial}{\partial r}\right]\vartheta_3^2 \psi\right\} = 0, \tag{63}$$

also on substituting from (61) and (63) in the fourth and fifth of (23) we have

$$\frac{1}{r}\frac{\partial}{\partial r}\vartheta_3^2 \phi = \frac{\partial^2}{\partial z^2}\left\{\vartheta_3^2 \phi + \left[\frac{\partial^2}{\partial r^2} - \frac{2}{r}\frac{\partial}{\partial r}\right]\vartheta_3^2 \psi\right\},$$

$$\frac{1}{r^2}\frac{\partial}{\partial z}\left(\frac{1}{r}\frac{\partial}{\partial r}\vartheta_3^2 \psi - \vartheta_3^2 \phi\right) = \frac{\partial^2}{\partial z \partial r}\left\{\frac{1}{r}\vartheta_3^2 \phi + \frac{1}{r}\left[\frac{\partial^2}{\partial r^2} - \frac{2}{r}\frac{\partial}{\partial r}\right]\vartheta_3^2 \psi\right\}.$$

Simplifying the last two equations with the aid of (62), we have

$$\left.\begin{aligned} \frac{\partial}{\partial r}\vartheta_3^2 \phi &= \frac{\partial}{\partial r}\vartheta_3^2 \frac{\partial^2 \psi}{\partial z^2}, \\ \frac{\partial^2}{\partial r \partial z}\left\{\vartheta_3^2 \phi + \left[\frac{\partial^2}{\partial r^2} - \frac{3}{r}\frac{\partial}{\partial r}\right]\vartheta_3^2 \psi\right\} &= 0. \end{aligned}\right\} \tag{64}$$

22. The second of (64) combined with (62) requires that

$$\vartheta_3^2\phi + \left[\frac{\partial^2}{\partial r^2} - \frac{3}{r}\frac{\partial}{\partial r}\right]\vartheta_3^2\psi = 2Az + f(r), \tag{i}$$

A being an arbitrary constant and f denoting a function as yet indeterminate. Then the first of (64) requires that

$$\vartheta_3^4\psi = f(r) + F(z), \tag{ii}$$

F being also indeterminate; and this means that (i) may be replaced by

$$\vartheta_3^2\phi = \frac{\partial^2}{\partial z^2}\vartheta_3^2\psi + 2Az - F(z). \tag{iii}$$

By writing ψ for $\psi - \Phi(z) - \Psi(r)$, where

$$\Phi^{\text{iv}}(z) = F(z), \quad \left[\frac{d^2}{dr^2} - \frac{3}{r}\frac{d}{dr}\right]^2 \Psi(r) = f(r),$$

we shall not alter the forms of (60), (61), (62) or (64), and the relations (ii) and (iii) then simplify to

$$\vartheta_3^4\psi = 0 \quad \text{and} \quad \vartheta_3^2\phi = \frac{\partial^2}{\partial z^2}\vartheta_3^2\psi + 2Az, \quad \text{so that} \quad \vartheta_3^4\phi = 0. \tag{iv}$$

Then (63) can be replaced by

$$\frac{\partial}{\partial z}\left\{\frac{\Theta}{1+\sigma} + \frac{2Az}{r} + \frac{1}{r^2}\frac{\partial}{\partial r}\vartheta_3^2\psi\right\} = 0,$$

which, combined with the requirement $\nabla^2\Theta = 0$,* has the solution

$$\frac{\Theta}{1+\sigma} + \frac{2Az}{r} + \frac{1}{r^2}\frac{\partial}{\partial r}\vartheta_3^2\psi = Pr + \frac{Q}{r}, \tag{v}$$

P and Q being arbitrary constants.

We can simplify again by superposing on ψ a function of r only and a function of z only such that in our equations

$$\vartheta_3^2\psi \text{ must be replaced by } \vartheta_3^2\psi + \tfrac{1}{4}Pr^4 + \tfrac{1}{2}Qr^2 - \tfrac{1}{3}Az^3. \tag{vi}$$

Thereby (60) and (61) are left unaltered, since no alteration is entailed in $\frac{\partial^2\psi}{\partial r\,\partial z}$; equations (iv) are replaced by

$$\left.\begin{aligned} \vartheta_3^4\psi &= 2Q + 2Az, \\ \vartheta_3^2\phi &= \frac{\partial^2}{\partial z^2}\vartheta_3^2\psi, \\ \vartheta_3^4\phi &= 0; \end{aligned}\right\} \tag{65}$$

---

* ∇^2 being defined as in (24).

and equation (v) reduces to

$$\frac{\Theta}{1+\sigma}+\frac{2Az}{r}+\frac{1}{r^2}\frac{\partial}{\partial r}\vartheta_3^2\psi = 0. \tag{66}$$

23. Now, from the definition of Θ, we have

$$r^2(\widehat{rr}+\widehat{\theta\theta}) = -2(1+\sigma)Azr-\frac{\partial}{\partial r}\{(1+\sigma)\vartheta_3^2\psi+\phi\}. \tag{67}$$

Combining this with the first of (60) we have

$$r^2(2\widehat{rr}+\widehat{\theta\theta}+\widehat{r\theta}) = -2(1+\sigma)Azr-\frac{\partial}{\partial r}\{(1+\sigma)\vartheta_3^2\psi+\phi\}-\frac{\partial^3\psi}{\partial r\,\partial z^2},$$

and combining this again with the first of (22) we obtain

$$\begin{aligned}\frac{\partial}{\partial r}\left\{r^3.\widehat{rr}+\phi+(1+\sigma)\vartheta_3^2\psi+2\frac{\partial^2\psi}{\partial z^2}+(1+\sigma)Azr^2\right\}\\ = r\frac{\partial^2\phi}{\partial z^2} = r\vartheta_3^2\phi-\left[r\frac{\partial^2}{\partial r^2}-3\frac{\partial}{\partial r}\right]\phi\\ = r\frac{\partial^2}{\partial z^2}\vartheta_3^2\psi-\left[r\frac{\partial^2}{\partial r^2}-3\frac{\partial}{\partial r}\right]\phi, \quad \text{by (65)},\\ = r(2Q+2Az)-\left[r\frac{\partial^2}{\partial r^2}-3\frac{\partial}{\partial r}\right](\phi+\vartheta_3^2\psi), \quad \text{by (65) again.}\end{aligned}$$

Integrating, we deduce that

$$r^3.\widehat{rr} = (Q-\sigma Az)r^2-2\frac{\partial^2\psi}{\partial z^2}-\left[r\frac{\partial}{\partial r}-3\right]\phi-\left[r\frac{\partial}{\partial r}-(3-\sigma)\right]\vartheta_3^2\psi+F(z), \tag{68}$$

$F(z)$ denoting a *new* (undetermined) function of z; and then we have from (67)

$$\left.\begin{aligned}r^3.\widehat{\theta\theta} &= -\{Q+(2+\sigma)Az\}r^2+2\frac{\partial^2\psi}{\partial z^2}-3\phi-F(z)\\ &\qquad -\left[\sigma r\frac{\partial}{\partial r}+3-\sigma\right]\vartheta_3^2\psi,\\ &\text{and from (60)}\\ -r^3.\widehat{r\theta} &= (Q-\sigma Az)r^2+F(z)-\left[r\frac{\partial}{\partial r}-3\right]\phi+\left[r\frac{\partial}{\partial r}-2\right]\frac{\partial^2\psi}{\partial z^2}\\ &\qquad -\left[r\frac{\partial}{\partial r}-(3-\sigma)\right]\vartheta_3^2\psi.\\ \text{Hence}\quad r^3(\widehat{rr}-\widehat{\theta\theta}) &= 2(Q+Az)r^2+2F(z)-4\frac{\partial^2\psi}{\partial z^2}-\left[r\frac{\partial}{\partial r}-6\right]\phi\\ &\qquad -\left[(1-\sigma)r\frac{\partial}{\partial r}-2(3-\sigma)\right]\vartheta_3^2\psi.\end{aligned}\right\} \tag{69}$$

These expressions, with (61), as they stand satisfy all of (22) when account is taken of the relations (65); and they also satisfy all of (23) provided that

$$F''(z) = 2\sigma(Q - Az). \tag{70}$$

Stresses induced by load applied to a toroidal 'hook' (Problem (ii))

24. Case (ii) finds practical application in the design of hooks. Here (figure 2) the form of the elastic solid approximates to an incomplete tore merging into a cylindrical 'shank' (S) of which the axis (usually) coincides with the line of application of the load (W). The largest stresses occur in that cross-section (AB) which is perpendicular to the line of loading, and 'Saint Venant's principle'* permits us to assert that they depend but little on the precise manner in which the load is applied or on the precise form of the hook near its point (P) or shank. Consequently a sufficiently exact theory can be based on the assumptions (a) that the hook may be replaced by an incomplete tore having a uniform cross-section similar to AB, (b) that the tore is loaded by tangential stresses applied to those sections which contain the line of loading. This 'equivalent tore' is indicated by bold lines in figure 2.

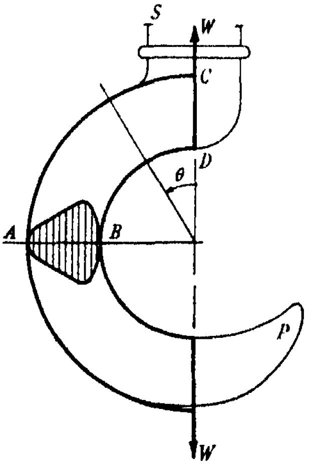


FIGURE 2

Clearly, if θ is measured from the upper loaded section CD, the assumptions stated in (21) are compatible with loading of the kind here contemplated. That is to say, the normal stress $\widehat{\theta\theta}$ will attain maximum intensity on AB and will be zero at the loaded sections, the reverse being true of $\widehat{r\theta}$ and $\widehat{\theta z}$. Expressions for $\widehat{r\theta}$, $\widehat{\theta z}$ from which the factor $\cos\theta$ is omitted give actual values of those stress components at the section $\theta = 0$, expressions for $\widehat{rr}$, $\widehat{\theta\theta}$, $\widehat{zz}$, $\widehat{zr}$ from which $\sin\theta$ is omitted give actual values of those stress components at the section $\theta = \frac{1}{2}\pi$.

25. The toroidal boundary of the hook is unstressed, so

$$\widehat{\nu r} = \widehat{\nu\theta} = \widehat{\nu z} = 0$$

in the relations (12), which (cf. § 7) are boundary conditions of our problem with the factors $\sin\theta$, $\cos\theta$ cancelled out. We now apply these conditions to define the undetermined functions which appear in (61), (68) and (69).

The first and second of (12), thus modified, show that

$$r(\widehat{rr} + \widehat{\theta\theta})\cos(r, \nu) + r(\widehat{zr} + \widehat{\theta z})\sin(r, \nu) = 0,$$

---

* Cf., for example, Southwell 1936, §§ 92–4.

whence
$$\frac{\partial}{\partial s}\left(\frac{1}{r}\frac{\partial^2\psi}{\partial r\,\partial z}\right) = 0,$$

when we substitute from (60). Therefore on the boundary

$$\frac{1}{r}\frac{\partial^2\psi}{\partial r\,\partial z} = \text{const.} = S \text{ (say)}. \tag{71}$$

From the last of (12) we have

$$\cos(r,\nu)\frac{1}{r}\frac{\partial^2\psi}{\partial z\,\partial r} + \frac{\partial\phi}{\partial s} = 0,$$

when we substitute from (61), or

$$\frac{\partial\phi}{\partial s} = S\,.\frac{\partial z}{\partial s},$$

in virtue of (71). Therefore on the boundary

$$\phi = Sz + T \text{ (say)}. \tag{72}$$

The first of (12), written in the form

$$r^2\{\widehat{rr}\cos(r,\nu) + \widehat{zr}\sin(r,\nu)\} = 0,$$

requires that

$$\cos(r,\nu)\left\{Qr - \sigma Azr - \frac{2}{r}\frac{\partial^2\psi}{\partial z^2} + \frac{F(z)}{r} - \left[\frac{\partial}{\partial r} - \frac{3}{r}\right]\phi \right.$$
$$\left. - \left[\frac{\partial}{\partial r} - \frac{3-\sigma}{r}\right]\vartheta_3^2\psi\right\} + \sin(r,\nu)\left(\frac{1}{r}\frac{\partial^2\psi}{\partial z\,\partial r} - \frac{\partial\phi}{\partial z}\right) = 0. \tag{73}$$

26. We shall satisfy the boundary conditions (71) and (72), and without contributing to any of the stress components defined by (61), if we make

$$\phi = Sz + T, \quad \frac{\partial^2\psi}{\partial z\,\partial r} = Sr, \tag{i}$$

everywhere; and thereby we shall make no contribution to the other stress components as given by (68) and (69) with

$$F(z) = \sigma(Qz^2 - \tfrac{1}{3}Az^3) \tag{74}$$

(as is consistent with (70)), if at the same time we make ψ satisfy the conditions

$$\frac{\partial}{\partial r}\vartheta_3^2\psi = 0, \quad 3(Sz+T) + (3-\sigma)\vartheta_3^2\psi + L(z) - 2\frac{\partial^2\psi}{\partial z^2} = 0, \tag{ii}$$

in which $L(z)$ stands for the linear part of $F(z)$ as obtained by integration of (70).

Now the general solution of the second of (i) is

$$\left.\begin{aligned} \psi &= \tfrac{1}{2}Szr^2 + (\text{function of } r \text{ only}) + (\text{function of } z \text{ only}) \\ &= \tfrac{1}{2}Szr^2 + f_1(r) + f_2(z), \text{ say,} \\ \text{and then}\quad \frac{\partial^2\psi}{\partial z^2} &= f_2''(z), \\ \vartheta_3^2\psi &= -2Sz + r^3\frac{d}{dr}\frac{1}{r^3}f_1'(r) + f_2''(z). \end{aligned}\right\} \tag{iii}$$

So we shall satisfy the first of (ii), together with the second and third of (65), provided that

$$r^3\frac{d}{dr}\frac{1}{r^3}\frac{d}{dr}f_1(r) = \text{const.},\quad f_2^{\text{iv}}(z) = 0. \tag{iv}$$

Evidently (iv) are compatible with the second of (ii): so we incur no loss of generality if, retaining (61) and (65) unaltered, we replace (68) and (69) by

$$\left.\begin{aligned} r^3\,.\,\widehat{rr} &= Q(r^2+\sigma z^2) - \sigma A(zr^2 + \tfrac{1}{3}z^3) - \left[r\frac{\partial}{\partial r} - 3\right]\phi - 2\frac{\partial^2\psi}{\partial z^2} \\ &\qquad - \left[r\frac{\partial}{\partial r} - (3-\sigma)\right]\vartheta_3^2\psi, \\ r^3\,.\,\widehat{\theta\theta} &= -Q(r^2+\sigma z^2) - A\{(2+\sigma)zr^2 - \tfrac{1}{3}\sigma z^3\} - 3\phi + 2\frac{\partial^2\psi}{\partial z^2} \\ &\qquad - \left[\sigma r\frac{\partial}{\partial r} + (3-\sigma)\right]\vartheta_3^2\psi, \\ -r^3\,.\,\widehat{r\theta} &= Q(r^2+\sigma z^2) - \sigma A(zr^2 + \tfrac{1}{3}z^3) - \left[r\frac{\partial}{\partial r} - 3\right]\phi \\ &\qquad + \left[r\frac{\partial}{\partial r} - 2\right]\frac{\partial^2\psi}{\partial z^2} - \left[r\frac{\partial}{\partial r} - (3-\sigma)\right]\vartheta_3^2\psi, \end{aligned}\right\} \tag{75}$$

and replace the boundary conditions (71)–(73) by

$$\left.\begin{aligned} &\phi = \frac{\partial^2\psi}{\partial z\,\partial r} = 0, \\ &\cos(r,\nu)\Big\{Q(r^2+\sigma z^2) - \sigma A(zr^2 + \tfrac{1}{3}z^3) - 2\frac{\partial^2\psi}{\partial z^2} \\ &\qquad - \left[r\frac{\partial}{\partial r} - (3-\sigma)\right]\vartheta_3^2\psi\Big\} = r\frac{\partial\phi}{\partial\nu}. \end{aligned}\right\} \tag{76}$$

27. Equations (61), (65), (75) and (76) constitute our formal solution of the hook problem. For equilibrium of the hook as a whole we must have

$$W = -\iint \widehat{r\theta}\,dr\,dz = \iint \widehat{\theta\theta}\,dr\,dz, \tag{77}$$

when $\widehat{r\theta}$, $\widehat{\theta z}$ are the quantities so denoted in §§ 7, 20–23, and when the surface integrals extend to the whole area of the cross-section; and by considering the implications of (21) as applied to the entire half-tore of figure 2 it can be shown that the resultant of $\widehat{r\theta}$, $\widehat{\theta z}$ must be directed along Or. That these requirements are satisfied can be established by appeal to the second and third of (22), which (combined with the boundary conditions of § 25) show that

$$-\iint (\widehat{r\theta}+\widehat{\theta\theta})\,dr\,dz = \oint r\{\widehat{r\theta}\cos(r,\nu)+\widehat{\theta z}\sin(r,\nu)\}\,ds$$

$$= 0 \quad \text{in virtue of the second of (12),}$$

$$\iint \widehat{\theta z}\,dr\,dz = \iint\left\{\frac{\partial}{\partial r}(r\,.\,\widehat{rz})+\frac{\partial}{\partial z}(r\,.\,\widehat{zz})\right\}dr\,dz$$

$$= \oint r\{\widehat{zr}\cos(r,\nu)+\widehat{zz}\sin(r,\nu)\}\,ds$$

$$= 0 \quad \text{in virtue of the third of (12).}$$

We conclude that the only resultant action which is compatible with the assumption (21) is a tension applied as shear stress at those sections where $\widehat{rr}$, $\widehat{\theta\theta}$, $\widehat{zz}$, $\widehat{zr}$ vanish severally, i.e. two opposed forces W of magnitude given by (77) and directed perpendicularly to the axis of revolution. *But the point at which these forces cut Oz is not restricted*; i.e. in general they have moments about the origin of magnitude

$$\mathbf{N} = \iint (r\,.\,\widehat{\theta z} - z\,.\,\widehat{r\theta})\,dr\,dz,$$

if the moment axis is taken in each instance as having the direction of θ for the section considered, and if $\widehat{r\theta}$, $\widehat{\theta z}$ have the same significance as in §§ 7, 20–23. Thus two independent types of loading are covered by the assumption (21), and an explanation is forthcoming of the two independent constants (Q and A) in our solution.

III. COMPARISON WITH EARLIER SOLUTIONS. SIMPLE EXAMPLES

28. We conclude this paper with a brief review of our solutions from the historical standpoint.

Problem (i, 1) *a* (§§ 9–10)

This problem was investigated by Michell (1899), who gave exact solutions for a family of approximately circular cross-sections, also some discussion (not extending to quantitative results) of rectangular cross-sections. His paper indicates the practical application to helical springs of small pitch (cf. § 10). It starts from *displacements* in formulating the governing equations, and for approximately circular cross-sections it presents diagrams exhibiting the warping of cross-sections under load. Its formal solution covers aeolotropy of an axially symmetrical kind, becoming equivalent to that of §§ 9–10 in the special case of isotropic material.

The problem was also studied by O. Gohner (1930–1), with results which have been reviewed by Timoshenko (1934). In place of an exact solution for an approximation to the desired boundary, Gohner sought the solution for an exactly circular, elliptical or rectangular boundary by a process of successive approximation.

No simple solution (having any practical interest) seems to have been discovered.

Problem (i, 1)*b* (§ 11)

29. This problem too seems to have been first solved by Michell (1899), who derived it as a particular case of his solution for incomplete tores, but gave no quantitative results. Rediscovery of the solution by Föppl (1905) led to studies by Willers and others which have been reviewed by Love (1927, § 226 A) and by Timoshenko (1934).

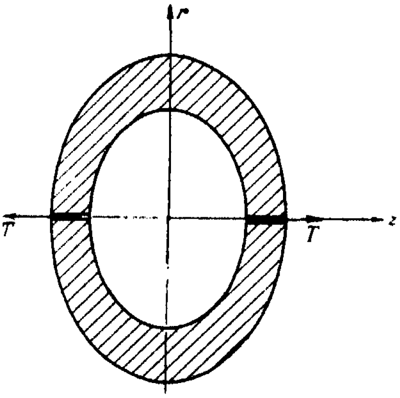


FIGURE 3

An obvious solution of equation (33) is

$$\phi = Ar^4 \quad (A \text{ constant}),$$

giving

$$\widehat{r\theta} = 0, \qquad \widehat{\theta z} = -4Ar, \tag{78}$$

according to (26). On the boundary, according to (30), r has a constant value; so (78) represents the familiar solution for a circular shaft of uniform diameter.

Simple solutions (but of little practical interest) can be obtained by the indirect method. Thus a stress function

$$\phi = r^2 + 2z^2$$

satisfies (33) and gives, according to (26),

$$\widehat{r\theta} = 4\frac{z}{r^2}, \qquad \widehat{\theta z} = -\frac{2}{r}. \tag{79}$$

According to (30) any spheroid generated by rotation about Oz of an ellipse of the family

$$r^2 + 2z^2 = \text{const.}$$

will be a possible boundary. The stress system (79) could be generated by suitable tractions applied to the surface of a small hole drilled concentrically with the axis of symmetry (figure 3).

Problem (i, 2)*a* (§§ 16–17)

30. This problem was solved by Love (1927, § 188) in a treatment starting from deflexions: his work is reproduced by Timoshenko (1934), who gives further references. Michell (1899) showed the problem to be a special case of the problem of flexure for incomplete tores (cf. § 16), but made no attempt to deal quantitatively with particular cases.

The formal solutions of Love and Michell differ both from one another and from the solution given in (52). Love's solution makes

$$\left.\begin{aligned} \widehat{rr}+\widehat{\theta\theta} &= \frac{\partial}{\partial z}\left\{\frac{\partial^2\chi}{\partial z^2} - (1-2\sigma)\nabla^2\chi\right\}, \\ \widehat{rr}-\widehat{\theta\theta} &= -\frac{\partial}{\partial z}\left(r\frac{\partial}{\partial r}\frac{1}{r}\frac{\partial\chi}{\partial r}\right), \\ \widehat{zz} &= \frac{\partial}{\partial z}\left\{(2-\sigma)\nabla^2\chi - \frac{\partial^2\chi}{\partial z^2}\right\}, \\ \widehat{zr} &= \frac{\partial}{\partial r}\left\{(1-\sigma)\nabla^2\chi - \frac{\partial^2\chi}{\partial z^2}\right\}, \end{aligned}\right\} \quad (80)$$

where

$$\nabla^2 \equiv \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{\partial^2}{\partial z^2},$$

and χ is a stress function satisfying

$$\nabla^4\chi = 0.$$

These expressions become identical with (52) if

$$\left.\begin{aligned} \frac{1}{r}\frac{\partial}{\partial r}\vartheta^2\lambda &= \frac{\partial}{\partial z}\nabla^2\chi, \\ -\frac{\partial}{\partial z}\vartheta^2\lambda &= r\frac{\partial}{\partial r}\nabla^2\chi, \\ 2\sigma\vartheta^2\lambda - \frac{\partial^2\lambda}{\partial z^2} &= r\frac{\partial^2\chi}{\partial r\,\partial z}, \end{aligned}\right\} \quad (81)$$

and (81) are evidently compatible in virtue of the equations governing λ and χ. Michell's solution is a special case ($\beta = 0$) of his solution discussed below (§ 32).

31. The well-known expressions for the stresses induced by pressures applied to the surfaces of a uniform tube can be deduced from (52) by writing

$$(1+\sigma)\vartheta^2\lambda = Ar^2, \quad \frac{\partial^2\lambda}{\partial z^2} = Br^2 + C, \tag{82}$$

—these expressions being compatible both with (45) and with one another. For then we have

$$\left.\begin{aligned} \widehat{rr}+\widehat{\theta\theta} &= 2B, \\ \widehat{rr}-\widehat{\theta\theta} &= -2C/r^2, \\ \widehat{zr} = 0, \quad \widehat{zz} &= 2(A-B), \end{aligned}\right\} \tag{83}$$

and the constants A, B and C can be adjusted to satisfy the boundary conditions.

Another simple solution, in which all four of the stress components (52) are non-zero, is that which gives the effect of a small spherical hole upon an otherwise uniform tension. This was propounded in a paper (Gough & Southwell 1926) which attempted to explain the occurrence of 'spiral' fractures in crankshafts.

Problem (i, 2) *b* (§§ 18–19)

32. Michell's treatment of this problem started as before from displacements, and it terminated with a formal solution of the governing equations—no quantitative results were obtained. *In our notation* his stress function ψ satisfies the equation

$$\vartheta^4\psi = -\frac{2(\lambda+\mu)}{\lambda+2\mu}\beta \quad (\beta \text{ constant}) \tag{84}$$

and his expressions corresponding with (56) are

$$\left.\begin{aligned} \widehat{rr}+\widehat{\theta\theta} &= \frac{2\mu}{r}\frac{\partial}{\partial r}\left(\tfrac{1}{2}\beta r^2 - \vartheta^2\psi - \frac{\partial^2\psi}{\partial z^2}\right), \\ \widehat{rr}-\widehat{\theta\theta} &= -2\mu\beta - \frac{2\mu}{\lambda+\mu}\left[\frac{1}{r}\frac{\partial}{\partial r} - \frac{2}{r^2}\right]\left\{\mu\vartheta^2\psi + (\lambda+\mu)\frac{\partial^2\psi}{\partial z^2}\right\}, \\ \widehat{zz} &= \qquad -\frac{\mu}{\lambda+\mu}\frac{1}{r}\frac{\partial}{\partial r}\left\{\lambda\vartheta^2\psi - 2(\lambda+\mu)\frac{\partial^2\psi}{\partial z^2} - \tfrac{1}{2}\lambda\beta r^2\right\}, \\ \widehat{zr} &= \qquad \frac{\mu}{\lambda+\mu}\frac{1}{r}\frac{\partial}{\partial z}\left\{\lambda\vartheta^2\psi - 2(\lambda+\mu)\frac{\partial^2\psi}{\partial z^2} - \tfrac{1}{2}\lambda\beta r^2\right\}, \\ \Theta &= \frac{\mu(3\lambda+2\mu)}{\lambda+\mu}\left(\beta - \frac{1}{r}\frac{\partial}{\partial r}\vartheta^2\psi\right), \end{aligned}\right\} \tag{85}$$

λ and μ being Lame's elastic constants. These are reconcilable with (56) provided that

$$\left.\begin{aligned} -\frac{\partial^2 \Omega}{\partial z^2} &= 2\mu\left(\vartheta^2\psi + \frac{\partial^2\psi}{\partial z^2}\right) + B\left\{\frac{\lambda}{2(\lambda+\mu)}r^2 - \frac{3\lambda+2\mu}{\lambda+2\mu}z^2 - \lambda C\right\}, \\ -\vartheta^2\Omega &= 2\mu\vartheta^2\psi + B\left\{\frac{\lambda}{2(\lambda+\mu)}r^2 - \frac{2(\lambda+\mu)}{\lambda+2\mu}z^2 - (\lambda+\mu)\,C\right\}, \\ -\sigma B &= \mu\beta, \end{aligned}\right\} \qquad (86)$$

where C is constant, B is the constant in equation (57), and Ω (to obviate confusion with Michell's use of μ) is substituted for the function μ of (56)–(57). The relations (86) are compatible with one another and with (57) and (84), so Michell's solution and ours can be identified.

33. A simple case of the general solution (56) is obtained by assuming conditions of *plane stress*, viz.

$$\widehat{zz} = \widehat{zr} = 0, \quad \text{everywhere.}$$

Then, according to (56), we have

$$\frac{\partial^2\mu}{\partial z^2} = (1+\sigma)\,\vartheta^2\mu + \text{const.}, \qquad \text{(i)}$$

and combining this equation with (57) we deduce that

$$\text{therefore} \qquad \left.\begin{aligned} \frac{\partial^2}{\partial z^2}\vartheta^2\mu &= 2(1+\sigma)\,B, \\ r\frac{\partial}{\partial r}\frac{1}{r}\frac{\partial}{\partial r}\vartheta^2\mu &= -2\sigma B, \end{aligned}\right| \qquad \text{(ii)}$$

by (57) again. The most general solution of (ii) is

$$\vartheta^2\mu = z(Cr^2+D) + Er^2 + F + B\{(1+\sigma)\,z^2 - \tfrac{1}{2}\sigma r^2(2\log r - 1)\}, \qquad \text{(iii)}$$

C, D, E, F being constants of integration. This can be combined with (i) to determine the form of μ.

The last step is not necessary to the finding of expressions for the stress components; for on substituting in (56) from (i) and (iii) we have

$$\left.\begin{aligned} \widehat{rr} + \widehat{\theta\theta} &= 2(1+\sigma)\,(Cz + E - \sigma B\log r), \\ \widehat{rr} - \widehat{\theta\theta} &= \sigma(1+\sigma)\,B + 2\sigma^2 B\frac{z^2}{r^2} - \frac{2}{r^2}\{\text{const.} + (1-\sigma)\,(zD+F)\}, \\ \widehat{zz} &= \widehat{zr} = 0. \end{aligned}\right\} \qquad \text{(iv)}$$

Neglecting the terms in z, and adjusting the constants so as to make $\widehat{rr}$ vanish at two radii a and b, we deduce for the stresses in a *narrow* rectangular section the expressions

$$\left.\begin{aligned} \widehat{rr} &= -\frac{\sigma(1+\sigma)\,B}{b^2-a^2}\left\{\frac{a^2b^2}{r^2}\log\frac{b}{a}+b^2\log\frac{r}{b}+a^2\log\frac{a}{r}\right\}, \\ \widehat{\theta\theta} &= \frac{\sigma(1+\sigma)\,B}{b^2-a^2}\left\{\frac{a^2b^2}{r^2}\log\frac{b}{a}-b^2\log\frac{r}{b}-a^2\log\frac{a}{r}-(b^2-a^2)\right\}, \end{aligned}\right\} \tag{87}$$

which agree with the solution of Golovin (1881) as quoted by Timoshenko (1934, § 23). This special case of Problem (i, 2) b has also been solved by Ribiere (1889) and Guest (1918).

Timoshenko (1934, § 112) also summarizes a treatment by Gohner of Problem (i, 2) b as relating to circular cross-sections. The method adopted was one of successive approximation (cf. § 28).

Problem (ii) (§§ 24–27)

34. It does not appear that any completely general treatment of this problem has been given previously. Timoshenko's account (1934, § 27) of work published in Russian by Golovin (1881) suggests that his investigation was restricted to hooks of *narrow rectangular* cross-section (a case of plane stress).

This case may be treated on the lines of § 33. Assuming that

$$\widehat{zz} = \widehat{zr} = 0, \quad \textit{everywhere}, \tag{i}$$

we deduce from (61) that ϕ is a function of z only, therefore must vanish everywhere in virtue of the first of (76). Then from (61) again we have

$$\frac{\partial^2\psi}{\partial z\,\partial r} = 0, \quad \textit{everywhere}, \tag{ii}$$

so that $\widehat{\theta z}$ also vanishes; and from the first and second of (65) we have

$$\frac{\partial^2}{\partial z^2}\vartheta_3^2\psi = 0, \quad r^3\frac{\partial}{\partial r}\frac{1}{r^3}\frac{\partial}{\partial r}\vartheta_3^2\psi = 2Q+2Az. \tag{iii}$$

Equation (ii) requires ψ to have the form

$$\left.\begin{aligned} \psi &= F_1(r)+F_2(z), \\ \vartheta_3^2\psi &= \left[r^3\frac{d}{dr}\frac{1}{r^3}\frac{d}{dr}\right]F_1(r)+F_2''(z), \end{aligned}\right\} \tag{iv}$$

so we have (the second line above)

where, in virtue of (iii),

$$F_2^{\text{iv}}(z) = 0, \quad \left[r^3\frac{d}{dr}\frac{1}{r^3}\frac{d}{dr}\right]^2F_1(r) = 2Q, \quad A = 0. \tag{v}$$

Then according to (iv)

$$\left.\begin{aligned}\frac{\partial^2\psi}{\partial z^2} &= F_2''(z) = E+Fz,\\ \vartheta_3^2\psi &= E+Fz-\tfrac{1}{2}Qr^2+G+Hr^4,\end{aligned}\right\}\tag{vi}$$

where E, F, G, H are constants of integration; and on substituting from (vi) in (75) *with the terms in A and ϕ suppressed* we have finally, as expressions for the stresses in a hook of *narrow* rectangular section (in which z has negligible values)

$$\left.\begin{aligned}\widehat{rr} &= 2Ar-\frac{2B}{r^3}+\frac{D}{r},\\ \widehat{\theta\theta} &= 6Ar+\frac{2B}{r^3}+\frac{D}{r},\\ -\widehat{r\theta} &= 2Ar-\frac{2B}{r^3}+\frac{D}{r},\end{aligned}\right|\text{ with } \widehat{zz}=\widehat{zr}=\widehat{z\theta}=0,\tag{88}$$

where

$$2A=-(1+\sigma)H,\quad 2B=-(1-\sigma)E-(3-\sigma)G,\quad 2D=(1+\sigma)Q.\tag{89}$$

These are Golovin's results as stated by Timoshenko (*loc. cit.*).

Grateful acknowledgment is due to Mr F. S. Shaw for his kindness in preparing the diagrams which accompany this paper.

REFERENCES

Föppl, A. 1905 *S.B. bayer. Akad. Wiss.* **35**, 249, 504.

Gohner, O. 1930–1. *Ingen. Arch.* **1**, 619 and **2**, 1, 381.

Golovin, H. 1881 *Trans. Inst. Tech., St Petersburg.* (Cited by Timoshenko, 1934, p. 60.)

Gough, H. J. & Southwell, R. V. 1926 *Phil. Mag.* **1**, 71–97.

Guest, J. J. 1918 *Proc. Roy. Soc.* A, **95**, 1–21.

Love, A. E. H. 1927 *The mathematical theory of elasticity*, 4th ed. Camb. Univ. Press.

Michell, J. H. 1899 *Proc. Lond. Math. Soc.* **31**, 130–146.

Ribiere, C. 1889 *C.R. Acad. Sci., Paris*, **108**, 561–3 (also **132**, 315–7 and *errata* 448).

Southwell, R. V. 1936 *Introduction to the theory of elasticity.* Oxford Univ. Press.

Timoshenko, S. P. 1934 *Theory of elasticity.* New York: McGraw-Hill Co.

Effect of thermal vibrations on the scattering of X-rays. III

BY M. BORN, F.R.S.
University of Edinburgh

(*Received* 2 *December* 1941)

The first section of this paper contains a comparison of Sir C. V. Raman's explanation of the diffuse scattering with the so-called 'thermal' theory. It is shown that Raman's demand of a theory based on quantum principles is satisfied by the modern forms of the thermal theory. The second part of the paper deals with the relation of the dynamical matrix (describing the mechanical properties of the lattice) and the scattering matrix (describing the scattering power). In the last section the influence of the ordinary Raman effect (mutual deformation of atoms) on the scattering is derived.

INTRODUCTION

As pointed out in Part I (Born & Sarginson 1941), the background scattering of X-rays is of fundamental importance for the kinetic theory of crystals, as it provides a method for determining, at least in principle, the dynamical properties of a lattice not in bulk (as the integrals representing specific heat and other such thermal properties) but in detail. This holds, of course, only if the experimental facts can really be explained by the thermal theory. There is a great amount of new experimental material, in the first instance contained in the article by Lonsdale & Smith (1941 *a*) published simultaneously with our Part I; their results seem to be in qualitative agreement with the thermal hypothesis. But Raman (1941), whose striking observations of the problem have contributed so much to the experimental side, rejects the thermal theory and insists on his own suggestion of a new quantum effect*.

The purpose of the present paper is a double one: I wish to remove the discrepancy of opinion by showing that it rests on a misunderstanding: Raman's demand that the explanation should be based on quantum principles is perfectly correct, but such a theory exists in fact and is nothing else than the modern form of the so-called 'thermal' theory which Raman rejects. In the rest of the paper I shall supplement the theory by a detailed consideration of the relation between crystal dynamics and scattering.

---

* [*Footnote added in proof.*] A full account of Raman's theory has now been published in a "Symposium" of papers on X-ray scattering (1941 *a*). It is surprisingly not based on quantum mechanics but quite on classical lines, and it introduces assumptions which are in contradiction to the dynamics of crystal lattices. In order to justify these Raman has published another "Symposium" of papers on the thermal energy of crystals (1941 *b*). A critical examination of these papers will appear elsewhere.

Raman's objections to the 'thermal' theory are partly based on experimental and partly on theoretical arguments. The experimental objections are due to the unfortunate fact that Raman has paid much attention to the behaviour of diamond, which cannot be called a 'normal' substance. As this point has been cleared up by Lonsdale & Smith (1941*b*), I shall restrict myself to a short remark: according to the investigations of Robertson, Fox & Martin (1934) there exist two types of diamond; the ordinary kind, type I, shows in its optical and electrical behaviour distinct traces of internal stress, while the rarer type II is a better approximation to an ideal lattice. According to Lonsdale & Smith, crystals of type I show anomalous types of extra spots which are essentially independent of temperature and presumably due to internal stress, while crystals of type II scatter in the normal way as expected on account of the thermal theory. The Indian investigators have apparently used crystals of the ordinary type I, in those orientations where the thermal spots are liable to be obscured by the stronger anomalous spots.

But Raman's main arguments are of a theoretical kind, as explained in a new paper by Raman & Nilakantan (1941). They maintain that the thermal agitation produces a rather uniform distribution of scattering power in the 'phase space', and they infer that this cannot give rise to the comparatively sharp maxima of intensity which they observe. Further, they remark 'that the intensity of the scattering in any specified direction being proportional to N (number of particles), its intensity would be quite negligible in relation to the intensity of the classical reflexion which is proportional to N^2'. Both these statements are incorrect as shown by the formulae derived in Part I. The background scattering power in the phase space is highly concentrated around the points of the reciprocal lattice (it has in fact an infinity in these points for the ideal case, namely, for exactly parallel X-rays [plane waves] and for a scattering domain large compared with the cell) and both Laue and background scattering are proportional to N.*

Though these objections of Raman & Nilakantan to the thermal theory are not valid, it is necessary to consider carefully their own suggestion of a new quantum effect. It seems to me that the dispute arises from the fact that the relevant papers on the thermal effect (Debye, Faxén, Waller) were written before the development of quantum mechanics. They are using quasi-classical language as in Bohr's quantum theory, and they assume that

---

* There is a misprint in Part I, formula (2·37); the periodic δ-function is $\delta = \lim_{n\to\infty} \delta_n$, with

$$\delta_n(Q) = \prod_{\alpha=1}^{3} \left(\frac{\sin \frac{1}{2} n Q_\alpha}{\sqrt{(2\pi n)} \sin \frac{1}{2} Q_\alpha} \right)^2;$$

its space integral over a region containing a lattice point $Q_\alpha = 2\pi\kappa_\alpha$ (κ_α integer) is unity, as for $|b-a| < \pi$

$$\lim_{n\to\infty} \frac{1}{n} \int_a^b \left(\frac{\sin nx}{\sin x}\right)^2 \phi(x)\,dx = \begin{cases} \pi\phi(0) & \text{if } a<0<b, \\ 0 & \text{if } a<0, b<0 \text{ or } a>0, b>0. \end{cases}$$

the frequency of the scattered X-rays is the same as that of the incident beam. This point has been corrected by v. Laue (1926) in a paper which Raman & Nilakantan quote as being 'the nearest approach to our point of view'; but they have overlooked that Laue explicitly states that his improvement involves no alteration whatever in Waller's and Faxén's results. This is in fact to be expected; for Laue's modification consists in taking account of the Doppler effect produced by the reflexion of the X-rays by the moving elastic waves, an effect which must be extremely small because of the smallness of the elastic quanta compared with the X-ray quanta. There are several more modern papers which systematically use wave mechanics. The most elaborate and rigorous of these is that of Ott (1935), who considers not only the vibrations of the lattice but also the electronic excitations of the atoms. Simpler methods are worked out by Zachariasen, Jauncey, and others (quoted in Part I), but they are all lacking in rigour at a deciding point. Our own work (Part I) was intended to improve this point and to work out the consequences of the theory in detail, which has independently been done also by Zachariasen (1941).

All these papers consider the scattering in exactly the same way as Raman & Nilakantan, namely, as a quantum interaction of the X-ray photons with the elastic quanta of the crystal lattices. They arrive, however, at the same formulae as the quasi-classical work of Waller and Faxén. I shall try to explain in § 1 of this paper why Raman's theory, though in the right direction, is not complete, as it considers only a part, in fact a rather small part, of the possible quantum interactions of X-ray and elastic waves. A complete consideration of all these quantum effects leads inevitably to the 'thermal' scattering formula.

In §§ 2 and 3 I shall treat the intimate relation between background scattering and crystal dynamics, and in the last section the modifications of the theory if mutual deformations of the atoms (Raman effect proper) are taken into account.

1. QUANTUM THEORY OF SCATTERING

Raman & Nilakantan have given only very short and qualitative descriptions of their theory*. I quote a sentence from their article in *Current Science* (1941): 'Basing ourselves on the accepted principles of the quantum theory of radiation, we showed that the lattice planes in a crystal should be capable of giving two kinds of geometric reflexion of X-rays; besides the classical or Laue reflexions, modified or quantum reflexions are also possible which have their origin in the *quantum-mechanical* excitation of the *optical* vibrations of the crystal lattice.'

In order to compare this assumption with the 'thermal' theory, it is necessary to interpret it in terms of the general scattering law of quantum mechanics which holds for any system carrying electric charges.

---

* See footnote on p. 397.

Let ν be the frequency of the incident light, $E_1, E_2, \ldots, E_r, \ldots$ the energy levels of the scattering system, r and s two states, r being not lower than s $(E_r \geqslant E_s)$, $\nu_{rs} = (E_r - E_s)/h$ the corresponding frequency, $\mathscr{M}^{(1)}_{rs}$ the matrix element of the electric moment of the system induced by the field $\mathscr{E}$ of the incident wave.

Then these two states give rise to two scattered waves (which in the case $E_r = E_s$ coincide), a 'normal' wave (N) with decreased frequency $\nu - \nu_{rs}$ (the system is lifted from the lower state s to the higher state r) and an 'anti-Stokes' wave (A) with increased frequency $\nu + \nu_{rs}$ (the system falls from the higher state r down to the lower one s). The intensities are (neglecting polarization effects)

$$\left.\begin{aligned} \mathscr{I}^N_{rs} &= \frac{64\pi^4}{3c^3}(\nu - \nu_{rs})^4 \, |\mathscr{M}^{(1)}_{rs}|^2 e^{-E_s/kT}, \\ \mathscr{I}^A_{rs} &= \frac{64\pi^4}{3c^3}(\nu + \nu_{rs})^4 \, |\mathscr{M}^{(1)}_{sr}|^2 e^{-E_r/kT}. \end{aligned}\right\} \tag{1·1}$$

The last (Boltzmann) factors account for the numbers of the atoms in the initial state.

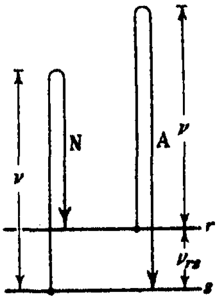


FIGURE 1

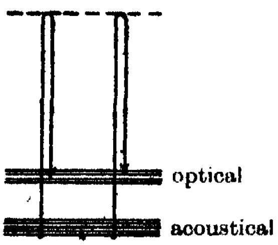


FIGURE 2

$\mathscr{M}^{(1)}_{rs}$ can be expressed in terms of the electric field of the incident wave $\mathscr{E}$ and of the electric moment $\mathscr{M}^0_{rs}$ of the unperturbed system, with the help of the generalized dispersion formula

$$\mathscr{M}^{(1)}_{rs} = \frac{1}{2h}\sum_k \left(\frac{\mathscr{M}^0_{rk}(\mathscr{M}^0_{ks}\mathscr{E})}{\nu_{ks} + \nu} - \frac{(\mathscr{M}^0_{rk}\mathscr{E})\mathscr{M}^0_{ks}}{\nu_{rk} + \nu} \right), \tag{1·2}$$

the summation k extended over all (virtual) states of the system. In forming the matrix elements the retardation of the waves, on their way from the source to the different parts of the system and from there to the point of observation, have to be taken into account. If the states r and s coincide one has the ordinary dispersion formula which gives rise to coherent scattering (without change of frequency). These formulae are completely general and must contain all possible scattering effects. Theories of a special phenomenon can only differ by the kind of approximation used. This depends mainly on the spacing of the energy levels of the system.

Consider now Raman's suggestion from this standpoint. For this purpose, a figure representing the energy levels of a crystal is drawn. There are electronic excitations possible, and as the X-ray quantum $h\nu$ is large compared with most of these levels they will be excited; they give rise to excitation waves ('excitons') and have been mentioned in our Part I, but as Raman does not consider them it will not be necessary here either. The energy spectrum consists then of the vibrational levels, and these form, for crystals with more than one atom in the cell, several groups: the acoustical branches ranging from $E = 0$ upwards to a certain maximum, and the optical branches, separated from the acoustical ones by a finite (but generally small) gap. Raman's assumption means that the background scattering is due to transitions from the ground level to the levels of the optical branches, as indicated in figure 2.

It is obvious that these transitions must occur. If they were the only ones excited by the incident X-ray the scattered intensity would be rigorously independent of temperature. I do not think that this assumption is made by Raman & Nilakantan, as they themselves have studied the temperature dependence of the new spots.

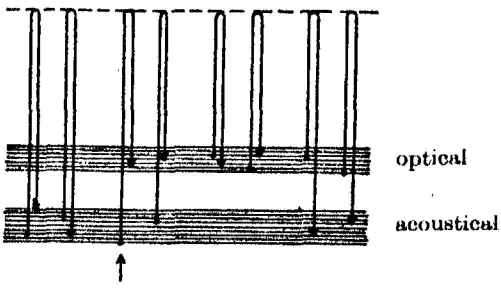


FIGURE 3

As a matter of fact, the probability of finding the crystal in the ground state at any finite temperature is extremely small. There is always a thermal excitation not only of all the levels of the acoustical branches, but also in a lesser degree of the optical branches. Hence the possible transitions will start from any of the vibrational levels, as indicated in figure 3, with a probability given by the product of $|\mathscr{M}^{(1)}_{rs}|^2$ with the corresponding Boltzmann factor. Now in the expression (1·2) for $\mathscr{M}^{(1)}_{rs}$ the summation over k extends over all virtual states (electronic transitions included), therefore its value is finite for all vibrational transitions (not only for $|r-s| = 0, \pm 2$ as it would be for purely vibrational levels), and $|\mathscr{M}^{(1)}_{rs}|$ can be replaced, according to Placzek, by the (rs)-matrix element of $\alpha\mathscr{E}$, where α is the polarizability of the system. Raman's transitions (figure 2), one of which is marked in figure 3 by an arrow, are only a small portion of the whole set. They are indeed those for which the energy of the lattice waves is completely taken from that of the incident X-ray. But all the other scattering transitions are at least of the same weight for the result.

The total scattered intensity is obtained by adding all normal and anti-Stokes intensities, given by (1·1) for any pair of levels of the vibrational spectrum, irrespective of the branch. This can be written as a double sum over r, s namely,

$$\mathscr{I} = \mathscr{I}_0 C \sum_r \sum_s |f(r,s)|^2 W_r, \quad W_r = e^{-E_r/kT}. \tag{1·3}$$

Here the vibrational frequencies ν_{rs} are neglected compared with the ν of the incident ray, and the factor $(\nu \pm \nu_{rs})^4 \sim \nu^4$ is combined with other factors (representing the scattering by a free electron) into the coefficient C; $f(r,s)$ represents the polarizability in terms of that of a free electron.

This formula (1·3) is strictly based on quantum theory; it contains Laue scattering and background scattering as well, as has been shown in Part I. But it must be admitted that this root of the 'thermal' theory is not easily recognizable in earlier publications and may be misunderstood.

In order to work out the matrix elements $f(r,s)$ for a spatially extended system (molecule, crystal) the retardations of the incident and scattered waves have to be taken into account. This leads to formula (1·2), Part I,

$$f = \sum_k f_k(\mathbf{u})\, e^{i(\mathbf{K}-\mathbf{K}')\cdot(\mathbf{r}_k+\mathbf{u}_k)}, \tag{1·4}$$

where the sum is extended over all atoms. This expression depends on the vibrations in virtue of the thermal displacements $\mathbf{u}_k$; these appear explicitly in the exponent, but also as arguments of the atomic scattering factors f_k, if the Raman effect is taken into account; for this means that the polarizability of each atom is influenced by the instantaneous positions of the other atoms, in first approximation linearly:

$$f_k = f_k^0 + \sum_{k'} \mathbf{f}_{kk'}^{(1)} \cdot \mathbf{u}_{k'}. \tag{1·5}$$

This effect was neglected in Part I, but will be discussed in § 4 of this paper.

The discussion of (1·4) leads automatically to the Laue formula as first approximation and an additional background term as second approximation. I shall discuss now in a simple way one of the points mentioned in the introduction which are used by Raman & Nilakantan as objections against the thermal theory.

If the displacements $\mathbf{u}_k$ of the atoms are expressed in terms of the normal co-ordinates ξ, the sum in (1·4) consists of terms each of which is a product of factors of the form $e^{i\mu\xi}$, where μ is the scalar product of $\mathbf{K}-\mathbf{K}'$ and a normalized eigen vector of the vibrations. One has therefore to calculate the matrix element $(e^{i\mu\xi})_{rs}$ for the transitions (r,s) of the oscillator ξ, and then the thermal average indicated in (1·3); it has been shown by Ott that this can (and must) be done rigorously, and I have given a new derivation of Ott's formula in the Appendix of Part I. The result is

$$\text{mean}\,(e^{i\mu\xi}) = e^{-\frac{1}{2}\mu^2\bar{\epsilon}/\omega^2}, \quad \bar{\epsilon} = \hbar\omega\left(\frac{1}{2}+\frac{1}{e^{\hbar\omega/kT}-1}\right) = \frac{\hbar\omega}{2}\coth\frac{\hbar\omega}{2kT}, \tag{1·6}$$

where $\bar{\epsilon}$ is the average energy of a linear oscillator, a function of the frequency ω which itself is a function of the wave vector of the elastic waves. Now the background scattering is essentially given by

$$1-e^{-\frac{1}{2}\mu\bar{\epsilon}/\omega^2},$$

a function which has a maximum for $\omega = 0$, and behaves, in some distance from this point, as $\frac{1}{2}\mu^2\bar{\epsilon}/\omega^2$. This shows that the scattering power is not due to the optical branches (as Raman & Nilakantan suggest), as for these ω does not approach 0, but to the acoustical branches, where $\omega \to 0$ for the points of the reciprocal lattice; hence there is a strong condensation of scattering power in the neighbourhood of the Laue spots which may produce relatively sharp maxima of the background (see Introduction).

The optical branches have, as is well known from the diatomic linear lattice, minima of ω in the middle of the reciprocal cell; it may therefore happen that weak background spots of another type exist which are just intermediate between Laue spots. But I do not know whether the experiments indicate the existence of such maxima.

2. DYNAMICAL TENSOR AND SCATTERING TENSOR

In order to abbreviate the writing of dynamical formulae, the following notation is adopted: If $k = 1, 2, \ldots, s$ is the basic index distinguishing the different simple lattices which compose the lattice of a crystal, then we write instead of the combinations

$$\left.\begin{matrix} 1x & 1y & 1z & 2x & 2y & 2z & \ldots & sx & sy & sz, \\ 1 & 2 & 3 & 4 & 5 & 6 & \ldots & (3s-2) & (3s-1) & 3s, \end{matrix}\right\} \tag{2·1}$$

and denote this index by α or β, etc. Any component of the displacement of a particle multiplied by the square root of its mass, $\sqrt{(m_k)}\,.\,u^l_{kx}, \ldots$, can then be written as v^l_α, and the second-order terms of the potential energy I (2·2)

$$\mathcal{U}_2 = \frac{1}{2}\sum_{ll'}\sum_{\alpha\beta} D^{l-l'}_{\alpha\beta}\, v^l_\alpha v^{l'}_\beta, \tag{2·2}$$

where $D^l_{\alpha\beta}$ is a new notation:

$$\sqrt{(m_k)}\,.\,u^l_{kx} \to v^l_\alpha, \quad \frac{1}{\sqrt{(m_k m_{k'})}}\begin{bmatrix} l \\ kk' \\ xy \end{bmatrix} \to D^l_{\alpha\beta}. \tag{2·3}$$

For each l the $D^l_{\alpha\beta}$ form a tensor in the $3s$-dimensional space, represented by the matrix

$$D^l = (D^l_{\alpha\beta}); \tag{2·4}$$

this is called the *dynamical tensor* for the cell l.

The formula I (2·7) can now be written

$$D_{\alpha\beta}(q) = \sum_l D^l_{\alpha\beta} e^{-i(lq)}, \tag{2·5}$$

and this quantity represents the dynamical tensor for a wave with wave vector $q(q_1, q_2, q_3)$, i.e. for a point in the reciprocal (or phase) space.

The linear equations for harmonic waves I (2·6) can now be written‡

$$\omega_j^2(q)\, e^j(q) - D(q)\, e^j(q) = 0, \tag{2·6}$$

where $e^j(q)$ is a column vector (components $e^j_\alpha(q)$, $\alpha = 1, 2, \ldots, 3s$).

Now there is a well-known algebraical theorem (see Appendix I) according to which one has for all positive or negative integers

$$\sum_j \omega_j^{2n}(q)\, e^j(q)\, e^{\dagger j}(q) = D^n(q), \tag{2·7}$$

where $e^{\dagger j}(q)$ means the conjugate transposed of $e^j(q)$, a row vector; or explicitly

$$\sum_j \omega_j^{2n}(q)\, e^j_\alpha(q)\, e^{*j}_\beta(q) = [D^n(q)]_{\alpha\beta}. \tag{2·7a}$$

If $f(\omega^2)$ is a function of ω^2 which can be expanded in (positive or negative) powers of ω^2,

$$f(\omega^2) = \sum_n c_n \omega^{2n}, \tag{2·8}$$

one has

$$\sum_j f(\omega_j^2(q))\, e^j(q)\, e^{\dagger j}(q) = f(D(q)). \tag{2·9}$$

The formula for the background scattering I (2·41) can now be written in the following form:

$$\mathscr{I}_B = \mathscr{I}_0 CN \sum_{kk'} \mathscr{A}_{kk'} G_k G^*_{k'}, \tag{2·10}$$

where

$$G_k = g_k e^{i(\mathbf{K}-\mathbf{K}').\mathbf{r}_k} = f_k e^{-U_k + i(\mathbf{K}-\mathbf{K}').\mathbf{r}_k}, \tag{2·11}$$

$$\mathscr{A}_{kk'} = \mathscr{B}_{kk'} - i\mathscr{C}_{kk'} = \sum_j \mathscr{L}^j_k(Q)\, \mathscr{L}^{*j}_{k'}(Q)\, \phi(\omega_j^2(Q)), \tag{2·12}$$

$$\mathscr{L}^j_k(q) = \frac{1}{\sqrt{m_k}} (\mathbf{K} - \mathbf{K}') . \mathbf{e}^j_k(q), \tag{2·13}$$

$$\phi(\omega^2) = \frac{\hbar^2}{kT} \frac{1}{2\beta} \coth \tfrac{1}{2}\beta, \quad \beta = \frac{\hbar\omega}{kT}, \tag{2·14}$$

$$Q_\alpha = (\mathbf{K} - \mathbf{K}') . \mathbf{a}_\alpha \quad (\alpha = 1, 2, 3). \tag{2·15}$$

---

‡ The factor m_k of ω_j^2, I (2·6), is wrong.

Solving the last equation with respect to $\mathbf{K}-\mathbf{K}'$ it follows that

$$\mathbf{K}-\mathbf{K}' = \sum_{\alpha} b_{\alpha} Q_{\alpha} = \mathbf{Q}, \tag{2·16}$$

$$(\mathbf{K}-\mathbf{K}') \,.\, \mathbf{r}_k = \sum_{\alpha} \beta_{\alpha}^{k} Q_{\alpha} = (\mathbf{r}_k \,.\, \mathbf{Q}), \quad \beta_{\alpha}^{k} = \mathbf{b}_{\alpha} \,.\, \mathbf{r}_k. \tag{2·17}$$

This shows that all quantities occurring in $\mathscr{I}_B$ are functions of Q_{α} which by (2·15) can be expressed in terms of the wave vectors of the incident and scattered wave. We can therefore write $\mathscr{I}_B$ as a function of the wave vector q_{α} of the elastic wave with the additional condition that q_{α} has everywhere to be replaced by $(\mathbf{K}-\mathbf{K}') \,.\, \mathbf{a}_{\alpha}$, or that the conservation law for momentum

$$q_{\alpha} = (\mathbf{K}-\mathbf{K}') \,.\, \mathbf{a}_{\alpha} \tag{2·18}$$

is fulfilled.

Now (2·12) can be written in the form

$$\mathscr{A}_{kk'} = \frac{1}{\sqrt{(m_k m_{k'})}} \sum_{xy} \sum_{j} e_{kx}^{j}(q)\, e_{ky}^{*j}(q)\, q_x q_y \phi(\omega_j^2(q)),$$

where, according to (2·16), $q_x = q_1 b_{1x} + q_2 b_{2x} + q_3 b_{3x}, \ldots$; and (2·10) in the form

$$\mathscr{I}_B = \mathscr{I}_0 CN \sum_{kx} \sum_{k'y} \sum_{j} e_{kx}^{j}(q) \frac{G_k}{\sqrt{m_k}} q_x e_{k'y}^{*j}(q) \frac{G_{k'}^{*}}{\sqrt{m_{k'}}} q_y \phi(\omega_j^2(q)).$$

Introducing the notation (2·1) for the indices (kx) and writing

$$w_{\alpha}(q) = w_{kx}(q) = \frac{G_k}{\sqrt{m_k}} q_x = \frac{f_k}{\sqrt{m_k}} (b_{1x} q_1 + b_{2x} q_2 + b_{3x} q_3)\, e^{-U_k + i(\beta^k q)}, \tag{2·19}$$

then

$$\mathscr{I}_B = \mathscr{I}_0 CN \sum_{\alpha\beta} S_{\alpha\beta}(q)\, w_{\alpha}(q)\, w_{\beta}^{*}(q), \tag{2·20}$$

where

$$S_{\alpha\beta}(q) = \sum_{j} e_{\alpha}^{j}(q)\, e_{\beta}^{*j}(q)\, \phi(\omega_j^2(q)) \tag{2·21}$$

is the element of the *scattering tensor* $S(q)$. The formula (2·9) can now be applied to obtain

$$S(q) = \phi(D(q)) = \frac{\hbar}{2} D^{-\frac{1}{2}}(q) \coth \frac{\hbar D^{\frac{1}{2}}(q)}{2kT}, \tag{2·22}$$

a relation which permits the calculation of the scattering tensor directly from the dynamical tensor, without solving the dynamical problem, i.e. without determining $\omega_j(q)$ and $e_k^{j}(q)$.

Using the well-known expansion

$$\tfrac{1}{2}\beta \coth \tfrac{1}{2}\beta = 1 - \sum_{n=1}^{\infty} (-1)^n \frac{B_n}{(2n)!} \beta^{2n},$$

where B_n are the Bernoullian numbers, then

$$\begin{aligned} \phi(\omega^2) &= \frac{\hbar^2}{kT} \frac{1}{2\beta} \coth \tfrac{1}{2}\beta = \frac{\hbar^2}{kT} \left\{ \beta^{-2} - \sum_{n=1}^{\infty} (-1)^n \frac{B_n}{(2n)!} \beta^{2(n-1)} \right\} \\ &= kT \left\{ \frac{1}{\omega^2} - \sum_{n=1}^{\infty} (-1)^n \frac{B_n}{(2n)!} \left(\frac{\hbar}{kT} \right)^{2n} \omega^{2(n-1)} \right\}, \end{aligned} \tag{2·23}$$

hence
$$S(q) = kT\left\{D^{-1}(q) - \sum_{n=1}^{\infty}(-1)^n \frac{B_n}{(2n)!}\left(\frac{\hbar}{kT}\right)^{2n} D^{n-1}(q)\right\}. \tag{2·24}$$

This remarkable formula is not new, but already contained in Waller's brilliant dissertation (formula (47), p. 29). But as this work is not widely known it seemed to me desirable to derive the result in the frame of presentation of the theory given here. Waller has applied his formula to the special case of a monatomic cubic crystal and shown that for the neighbourhood of a point of the reciprocal lattice (q_α equal to a multiple of 2π) the first term of the bracket in (2·24) can be expressed with the help of the elastic constants of the crystal. This method has been independently used by Miss Sarginson (1941) in our Part II and by Jahn (1941 *a*, *b*) for explaining the observations made by Lonsdale & Smith of the diffuse scattering by cubic crystals. Subsequent observations of the diffuse scattering by the soft, elastically anisotropic cubic crystals of sodium and lithium have provided a much more rigorous confirmation of the formula (Lonsdale & Smith 1941 *b*).

I wish to show here that these results can easily be generalized for any crystal lattice.

In the neighbourhood of $q_\alpha = 0$ or any equivalent point $q_\alpha = 2\pi\kappa_\alpha$ we need to consider only the acoustical branches (see end of § 1) $j = 1, 2, 3$; for these the $e^j_k(q)$ are independent of k, say $e^j(q)$, and satisfy the equations
$$\omega_j^2 e^j(q) - D(q)\, e^j(q) = 0, \tag{2·25}$$
where $D(q)$ is a three-dimensional matrix
$$D(q) = \left|\begin{matrix} D_{xx}(q) & D_{xy}(q) & D_{xz}(q) \\ D_{yx}(q) & D_{yy}(q) & D_{yz}(q) \\ D_{zx}(q) & D_{zy}(q) & D_{zz}(q) \end{matrix}\right|. \tag{2·26}$$

Each element is a quadratic form of the components of the vector
$$q_1\mathbf{b}_1 + q_2\mathbf{b}_2 + q_3\mathbf{b}_3 = \mathbf{q},$$
namely (see Born 1923),
$$D_{xy}(q) = \frac{1}{\rho}\sum_{\bar{x}\bar{y}}[x\bar{x}\,|\,y\bar{y}]\, q_{\bar{x}} q_{\bar{y}}, \tag{2·27}$$

where ρ is the density and $[x\bar{x}\,|\,y\bar{y}]$ are the elastic constants; or explicit with Voigt's notation, as matrix equation
$$\begin{Bmatrix} D_{xx} \\ D_{yy} \\ D_{zz} \\ D_{yz} \\ D_{zx} \\ D_{xy} \end{Bmatrix} = \frac{1}{\rho}\begin{Bmatrix} c_{11} & c_{66} & c_{55} & 2c_{65} & 2c_{51} & 2c_{16} \\ c_{66} & c_{22} & c_{44} & 2c_{24} & 2c_{46} & 2c_{62} \\ c_{55} & c_{44} & c_{33} & 2c_{43} & 2c_{35} & 2c_{54} \\ c_{65} & c_{24} & c_{43} & c_{23}+c_{44} & c_{45}+c_{36} & c_{64}+c_{25} \\ c_{51} & c_{46} & c_{35} & c_{45}+c_{36} & c_{31}+c_{55} & c_{56}+c_{14} \\ c_{16} & c_{62} & c_{54} & c_{64}+c_{25} & c_{56}+c_{14} & c_{12}+c_{66} \end{Bmatrix} \cdot \begin{Bmatrix} q_x^2 \\ q_y^2 \\ q_z^2 \\ q_y q_z \\ q_z q_x \\ q_x q_y \end{Bmatrix}. \tag{2·28}$$

The intensity is then approximately

$$\mathscr{I}_B = \mathscr{I}_0 CN\,kT\frac{G^2}{m}4\pi^2\sum_{xy}D^{-1}_{xy}(q)\,\kappa_x\kappa_y, \tag{2·29}$$

where $$\kappa_x = \kappa_1 b_{1x} + \kappa_2 b_{2x} + \kappa_3 b_{3x}$$

and $$\frac{G}{\sqrt{m}} = \sum_k \frac{G_k}{\sqrt{m_k}}. \tag{2·30}$$

This formula reduces in the case of a cubic crystal to that used by Jahn and Sarginson.

3. DETERMINATION OF THE DYNAMICAL MATRIX FROM THE OBSERVATION OF SCATTERING

The following considerations are perhaps at present not of practical importance, but of great principal interest. They are concerned with the possibility of solving the inverse problem:

Assume that, by measuring the scattering power, it would be possible to obtain numerical values for the matrix $S(q)$. Then the dynamical matrix $D(q)$ follows from (2·22),

$$D(q) = \psi(S(q)), \tag{3·1}$$

where ψ is the inverse function of ϕ.

Two questions have to be considered: (1) how to determine $S(q)$ from measurements of the intensity $\mathscr{I}_B$, and (2) how to expand the inverse function ψ of ϕ into powers of its argument.

(1) In the quadric (2·20) representing the intensity, the variables $w_\alpha(q)$ and the coefficients $S_{\alpha\beta}(q)$ depend on q; but the latter are periodic in the reciprocal lattice while the $w_\alpha(q)$ are not. Taking a set of n equivalent points, i.e. such that their q_α values differ by multiples of 2π, the $S_{\alpha\beta}(q)$ have the same values for these, but not the $w_\alpha(q)$. The latter are known functions of q_α given by (2·19) and have therefore known values in the n points. Hence the number n of the equivalent points can be chosen in such a way that the measurement of $\mathscr{I}_B/\mathscr{I}_0$ in these points supplies as many equations of the form (2·20) as there are unknown values $S_{\alpha\beta}(q)$. This can be done for any point in the basis cell of the reciprocal lattice. The $S_{\alpha\beta}(q)$ can therefore be considered as functions which are in principle accessible to empirical determination.

(2) Putting $$\frac{kT}{\hbar^2}\phi = y, \quad \beta^2 = \left(\frac{\hbar\omega}{kT}\right)^2 = x, \tag{3·2}$$

then (2·14) becomes

$$y = \frac{1}{2\sqrt{x}}\coth\tfrac{1}{2}\sqrt{x} = \frac{1}{x} + \frac{1}{12} - \frac{x}{720} + \dots, \tag{3·3}$$

and if $x = f(y)$ is the inverse function, then

$$\omega^2 = \frac{k^2T^2}{\hbar^2} f\left(\frac{kT}{\hbar^2}\phi\right) = \psi(\phi), \tag{3·4}$$

and (3·1)
$$D(q) = \frac{k^2T^2}{\hbar^2} f\left(\frac{kT}{\hbar^2}S(q)\right). \tag{3·5}$$

The inverse of (3·3) can be expanded into a series of powers of y^{-1}, namely,

$$x = f(y) = \frac{2}{y}\left(\tfrac{1}{2} + [a_1 + \tfrac{1}{8}]\frac{1}{y} + a_2\frac{1}{y^2} + a_3\frac{1}{y^3} + \ldots\right), \tag{3·6}$$

where
$$a_1 = -\tfrac{1}{12}, \quad a_2 = \tfrac{1}{360}, \quad a_3 = \tfrac{1}{7560}, \quad \ldots. \tag{3·7}$$

The a_n satisfy a linear recursive formula (see Appendix II).

Now (3·5) is explicitly given by the expansion

$$D(q) = kTS^{-1}(q)\left\{1 + \frac{1}{12}\frac{\hbar^2}{kT}S^{-1}(q) + \frac{1}{180}\frac{\hbar^4}{k^2T^2}S^{-2}(q) + \frac{1}{3780}\frac{\hbar^6}{k^3T^3}S^{-3}(q) + \ldots\right\}. \tag{3·8}$$

The harmonic force coefficients are obtained by reversing (2·5),

$$D(q) = \sum_l D^l e^{-i(lq)} \rightarrow D^l = \frac{1}{(2\pi)^3}\int D(q)\, e^{i(lq)}\, dq. \tag{3·9}$$

Summarizing the result it can be said that: A sufficiently thorough photometric survey of a properly selected set of background photographs is in principle equivalent to an experimental determination of the harmonic terms of the potential energy of interaction between all the particles of a lattice. The background photographs determine the dynamics of the lattice in the same way as the Laue-Bragg photographs the geometry of the lattice.

4. INFLUENCE OF THE MUTUAL DEFORMATION OF THE ATOMS ON THE BACKGROUND SCATTERING (RAMAN EFFECT PROPER)

The incoherent scattering with change of frequency (1·1), (1·2) was predicted by theory (Smekal 1923; Heisenberg & Kramers 1925; Born, Heisenberg & Jordan 1926) before the experimental discovery of Raman & Krishnan (1928) and of Landsberg & Mandelstam (1928) of this effect for vibrating molecules. The expression 'Raman effect' is therefore generally and properly used for this special case, and I shall use it here in the same sense. This effect is not contained in the formulae of Part I (and of § 2 of this paper), but, according to Placzek, it can easily be accounted for by considering the polarizability as function of the atomic displacements, as indicated in (1·4), (1·5). However, it is necessary to discuss the assumption on which these formulae rest.

The polarizability of an atom in a crystal will in general be represented not by a scalar but by a tensor; and even if this reduces for zero temperature to a scalar, owing to the high symmetry of the lattice, it will become a tensor for finite temperature in virtue of the vibrations. But this tensor can be split into a completely symmetrical part, equivalent to a scalar (represented by a diagonal matrix) and a rest which is a tensor without trace (vanishing diagonal sum). This separation is always performed for the discussion of the optical Raman effect, and it is shown that the two parts differ with respect to scattering in the following way: The symmetric part does not change the state of polarization of the incident light, while the traceless part produces depolarization.

In the case of X-rays we are not interested in polarization effects. Therefore it will be permitted to omit the traceless part and consider the polarizability as a scalar. This leads to the formula (1·5), and there is no difficulty in generalizing the theory of Part I by introducing the additional terms in the formulae and performing the necessary calculations. But these can be almost completely avoided by a trick suggested to me by Miss Sarginson. Instead of (1·5), write

$$f_k = f_k^0 \exp\left[\sum_{k'} \boldsymbol{\beta}_{kk'} \,.\, \mathbf{u}_{k'}\right], \tag{4·1}$$

which for small $\mathbf{u}_k$ is approximately the same expression. The exponential function in (4·1) can now be combined with that in (1·4) and the whole change in the calculation consists in replacing $(\mathbf{K}-\mathbf{K}') \,.\, \mathbf{u}_k$ by the complex expression

$$(\mathbf{K}-\mathbf{K}') \,.\, \mathbf{u}_k - i\sum_{k'} \boldsymbol{\beta}_{kk'} \,.\, \mathbf{u}_{k'} = \sum_{k'} \boldsymbol{\mu}_{kk'} \,.\, \mathbf{u}_{k'},$$

where
$$\boldsymbol{\mu}_{kk'} = (\mathbf{K}-\mathbf{K}')\,\delta_{kk'} - i\,\boldsymbol{\beta}_{kk'}.$$

If the system is a lattice one has instead of the index k the set $(k, l) = (k l_1 l_2 l_3)$ and

$$f_k^l = f_k^0 \exp\left[\sum_{k'l'} \boldsymbol{\beta}_{kk'}^{l-l'} \,.\, \mathbf{u}_{k'}^{l'}\right]. \tag{4·2}$$

As there is no mutual deformation of the atoms for a translation of the whole system the exponent must vanish if all $\mathbf{u}_k^l$ are equal; hence

$$\sum_{k'l'} \boldsymbol{\beta}_{kk'}^{l'} = 0. \tag{4·3}$$

The formula I (2·19) has now to be modified by replacing the term

$$i(\mathbf{K}-\mathbf{K}') \,.\, \mathbf{u}_k^l$$

in the exponent by $i\sum_{k'l'} \boldsymbol{\mu}_{kk'}^{l-l'} \,.\, \mathbf{u}_{k'}^{l'}$, where

$$\boldsymbol{\mu}_{kk'}^l = (\mathbf{K}-\mathbf{K}')\,\delta_{kk'}^l - i\,\boldsymbol{\beta}_{kk'}^l, \quad \delta_{kk'}^l = \begin{cases} 1 \text{ for } k = k',\ l = 0, \\ 0 \text{ for all other values of } k, k', l. \end{cases} \tag{4·4}$$

If now normal co-ordinates are introduced one has in I (2·22) instead of $\mathscr{L}^j_k(q)\, e^{i(lq)}/\sqrt{N}$ in the exponent the expression

$$\sum_{k'l'} \frac{1}{\sqrt{m_k}} (\boldsymbol{\mu}^{l-l'}_{kk'} \cdot \mathbf{e}^j_{k'}(q)) \frac{e^{i(l'q)}}{\sqrt{N}} = \Lambda^j_k(q) \frac{e^{i(lq)}}{\sqrt{N}},$$

or, instead of $\mathscr{L}^j_k(q)$,

$$\Lambda^j_k(q) = \mathscr{L}^j_k(q) - i \sum_{k'} \frac{1}{\sqrt{m_{k'}}} (\boldsymbol{\beta}_{kk'}(q) \cdot \mathbf{e}^j_{k'}(q))$$

$$= \sum_{k'} \frac{1}{\sqrt{m_k}} [\{(\mathbf{K} - \mathbf{K}')\, \delta_{kk'} - i\, \boldsymbol{\beta}_{kk'}(q)\} \cdot \mathbf{e}^j_k(q)], \tag{4·5}$$

where

$$\boldsymbol{\beta}_{kk'}(q) = \sum_{l'} \boldsymbol{\beta}^{l'}_{kk'}\, e^{-i(l'q)}. \tag{4·6}$$

In virtue of (4·3) one has

$$\sum_{k'} \boldsymbol{\beta}_{kk'}(0) = 0. \tag{4·7}$$

All the following calculations in Part I are unchanged apart from replacing $\mathscr{L}^j_k$ by Λ^j_k, and this holds also for the final formulae I (2·41). In the equivalent expressions of the present paper $\mathscr{L}^j_k$ appears only in (2·12); this has therefore to be replaced by

$$\mathscr{A}_{kk'} = \sum_j \Lambda^j_k \Lambda^{*j}_{k'}\, \phi(\omega^2_j). \tag{4·8}$$

If now the same operations are performed as in § 2, the formula (2·20) is again obtained with a modified definition of the $w_\alpha(q)$; instead of (2·19) one has

$$w_\alpha(q) = w_{kx}(q) = \frac{1}{\sqrt{m_k}} \left\{ G_k q_x - i \sum_{k'} G_{k'} \beta_{k'kx}(q) \right\}. \tag{4·9}$$

The additional imaginary term is periodic in q.

In order to estimate its influence on the scattering the coupling coefficients $\boldsymbol{\beta}^l_{kk'}$ ought to be known. They could be obtained, in principle, from observations of the ordinary Raman effect of molecules built of the same atoms, or by theoretical calculations. But at present there is nothing known about them. One can, however, consider the question how the new terms behave in the neighbourhood of points $q_\alpha = 2\pi\kappa_\alpha$ (κ_α integer) of the reciprocal lattice, i.e. for directions of the scattered beam near to a Laue spot. Expanding (4·6) with respect to the distance $q_\alpha - 2\pi\kappa_\alpha$ from a lattice point one gets

$$\boldsymbol{\beta}_{kk'}(q) = \sum_{l'} \boldsymbol{\beta}^{l'}_{kk'} - i \sum_{l'} \boldsymbol{\beta}^{l'}_{kk'} \cdot (l'(q - 2\pi\kappa)) + \dots. \tag{4·10}$$

All terms apart from the first one vanish for the Laue spot, $q_\alpha = 2\pi\kappa_\alpha$. Therefore the influence of the additional terms depends on the magnitude of $\sum_{l'} \boldsymbol{\beta}^{l'}_{kk'}$ compared with that of the aperiodic term which for $q_\alpha \to 2\pi\kappa_\alpha$ approaches a finite value $2\pi G_k(b_x\kappa)$. Now it is obvious that $\boldsymbol{\beta}^{l'}_{kk'}$ vanishes for a pair of equal atoms $(k, 0)$ and (k', l'); for it is impossible to define a vector by the mutual influence of two equal particles.‡ For different atoms, however, this is possible (e.g. a vector from the smaller to the larger one) and $\boldsymbol{\beta}^{l'}_{kk'}$ may be different from zero. It follows that for monatomic lattices the mutual deformation of the atoms has no effect on the scattering, while for polyatomic lattices such an effect may exist.

If the $\boldsymbol{\beta}^{l'}_{kk'}$ were known all considerations of the previous section on the determination of the dynamical matrix from the observation of background scattering would be still valid.

Appendix I

Let $e^j (j = 1, 2, \ldots, N)$ be a set of complex vectors (one-column matrices with elements e^j_α, $\alpha = 1, 2, \ldots, N$) and $e^{\dagger j}$ their adjoints (one-row matrices with conjugate elements e^{*j}_α).

Assume now that they are hermitean ortho-normal

$$e^{\dagger j} e^{j'} = \delta_{jj'}, \quad \text{or} \quad \sum_\alpha e^{*j}_\alpha e^{j'}_\alpha = \delta_{jj'}. \tag{I. 1}$$

It follows that
$$\sum_j e^j e^{\dagger j} = 1, \quad \text{or} \quad \sum_j e^j_\alpha e^{\dagger j}_\beta = \delta_{\alpha\beta}. \tag{I. 2}$$

Let $\lambda_1, \lambda_2, \ldots, \lambda_N$ be a set of real numbers; then

$$\Lambda^{(n)} = \sum_j \lambda^n_j e^j e^{\dagger j}, \quad \text{or} \quad \Lambda^{(n)}_{\alpha\beta} = \sum_j \lambda^n_j e^j_\alpha e^{*j}_\beta \tag{I. 3}$$

is a hermitean matrix. If n is an integer ($n = 0, \pm 1, \pm 2, \ldots$), then

$$\Lambda^{(n)}\Lambda = \sum_{jj'} \lambda^n_j \lambda_{j'} e^j e^{\dagger j} e^{j'} e^{\dagger j'} = \sum_j \lambda^{n+1}_j e^j e^{\dagger j'} = \Lambda^{(n+1)},$$

where Λ is written instead of $\Lambda^{(1)}$. It follows that

$$\Lambda^{(n)} = \Lambda^n \quad \text{for} \quad n = 0, \pm 1, \pm 2, \ldots. \tag{I. 4}$$

Now let the λ_j, e^j be the latent roots and vectors (eigen values and eigen vectors) of the hermitean matrix a $(= a^\dagger)$, so that

$$ae^j = \lambda_j e^j. \tag{I. 5}$$

Then by multiplying from the right with $e^{\dagger j}$ and summing over j

$$a \sum_j e^j e^{\dagger j} = \sum_j \lambda_j e^j e^{\dagger j}$$

is obtained, or with (I. 2)
$$a = \Lambda; \tag{I. 6}$$

hence
$$\Lambda^{(n)} = a^n, \quad \text{or} \quad \sum_j \lambda^n_j e^j e^{\dagger j} = a^n \quad (n = 0, \pm 1, \pm 2, \ldots). \tag{I. 7}$$

This is the theorem used in the text.

---

‡ This follows also from (4·3) if k can assume only one value.

APPENDIX II

The function
$$y = \frac{1}{2\sqrt{x}} \coth \tfrac{1}{2}\sqrt{x} \tag{II. 1}$$
can be expanded in powers of x, beginning with $x^{-1} + \ldots$; hence the inverse function $x = f(y)$ can be expanded in powers of $v = y^{-1}$, beginning with $x = v + \ldots$. Now
$$v = y^{-1} = \frac{2\sqrt{x}}{\coth \frac{1}{2}\sqrt{x}} \tag{II. 2}$$
satisfies the differential equation[‡]
$$\frac{dv}{dx} = \frac{1}{2} + \frac{v}{2x} - \frac{v^2}{8x}. \tag{II. 3}$$

Considering now v as the independent variable and introducing instead of $x(v)$ the function
$$w(v) = x(v) + v - \tfrac{1}{4}v^2; \tag{II. 4}$$
then $w(v)$ satisfies the differential equation
$$\frac{dw}{dv} = 3 - \tfrac{1}{2}v + (\tfrac{1}{2}v^2 - 2v)\frac{1}{w}. \tag{II. 5}$$

As $x(v) = v + \ldots$ the expansion of $w(v)$ begins with $2v + \ldots$, hence
$$w = 2v(1 + a_1 v + a_2 v^2 + \ldots). \tag{II. 6}$$

Then
$$\frac{1}{w} = \frac{1}{2v}(1 + b_1 v + b_2 v^2 + \ldots), \tag{II. 7}$$

where
$$\left.\begin{aligned} b_1 &= -a_1, \\ b_2 &= -a_1 b_1 - a_2 = a_1^2 - a_2, \\ b_3 &= -a_1 b_2 - a_2 b_1 - a_3 = -a_1^3 + 2a_1 a_2 - a_3, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ b_n &= -a_1 b_{n-1} - a_2 b_{n-2} - \ldots - a_n. \end{aligned}\right| \tag{II. 8}$$

One can write
$$b_n = (-1)^n \begin{vmatrix} a_1 & 1 & 0 & 0 & 0 & \ldots & 0 \\ a_2 & a_1 & 1 & 0 & 0 & \ldots & 0 \\ a_3 & a_2 & a_1 & 1 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{n-1} & a_{n-2} & a_{n-3} & \ldots & \ldots & \ldots & 1 \\ a_n & a_{n-1} & a_{n-2} & \ldots & \ldots & \ldots & a_1 \end{vmatrix}. \tag{II. 9}$$

Substituting (II. 6) and (II. 7) in (II. 5) one obtains
$$a_1 = -\tfrac{1}{12}, \quad 2(n+1)\,a_n = \tfrac{1}{4}b_{n-1} - b_n \quad \text{for} \quad n = 2, 3, \ldots.$$

[‡] I have to thank Professor E. T. Whittaker for suggesting this method to determine the coefficients of the expansion.

Using the last formula (II. 8) this can be written

$$(2n+1)\,a_n = (a_1+\tfrac{1}{4})\,b_{n-1}+a_2b_{n-2}+\dots+a_{n-1}b_1.$$

The right-hand side differs from the expression $-b_n$ (II. 8) only in so far as a_1 is replaced by $a_1+\frac{1}{4}$ and a_n by 0. Performing these operations in (II. 9) one obtains

$$a_n = \frac{(-1)^{n-1}}{2n+1}\begin{vmatrix} a_1+\frac{1}{4} & 1 & 0 & 0 & \dots & 0 \\ a_2 & a_1 & 1 & 0 & \dots & 0 \\ a_3 & a_2 & a_1 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n-1} & a_{n-2} & a_{n-3} & \dots & \dots & 1 \\ 0 & a_{n-1} & a_{n-2} & \dots & \dots & a_1 \end{vmatrix}. \qquad \text{(II. 10)}$$

This is a recursive formula expressing a_n by $a_1, a_2, \dots, a_{n-1}$. One finds

$$a_1 = -\tfrac{1}{12}, \quad a_2 = \tfrac{1}{360}, \quad a_3 = \tfrac{1}{7560}, \quad \dots. \qquad \text{(II. 11)}$$

The expansion of $x(v)$ is obtained by substituting (II. 6) in (II. 4)

$$\begin{aligned} x(v) &= w(v)-v+\tfrac{1}{4}v^2 = 2v(\tfrac{1}{2}+[a_1+\tfrac{1}{8}]\,v+a_2v^2+a_3v^3+\dots) \\ &= v(1+\tfrac{1}{12}v+\tfrac{1}{180}v^2+\tfrac{1}{3780}v^3+\dots). \end{aligned} \qquad \text{(II. 12)}$$

References

Born, M. 1923 *Atomtheorie des Festen Zustandes.* Leipzig: Teubner.

Born, M. & Sarginson, K. 1941 *Proc. Roy. Soc.* A, **179**, 69 (quoted as Part I).

Born, M., Heisenberg, W. & Jordan, P. 1926 *Z. Phys.* **35**, 557.

Heisenberg, W. & Kramers, H. A. 1925 *Z. Phys.* **31**, 681.

Jahn, H. A. 1941*a* *Nature, Lond.*, **147**, 511.

Jahn, H. A. 1941*b* *Proc. Roy. Soc.* A, **179**, 320.

Landsberg, G. & Mandelstam, L. 1928 *Naturwissenschaften*, **16**, 557, 772.

Laue, M. von 1926 *Ann. Phys., Lpz.*, **81**, 877.

Lonsdale, K. & Smith, H. 1941*a* *Proc. Roy. Soc.* A, **179**, 8.

Lonsdale, K. & Smith, H. 1941*b* *Nature, Lond.*, **148**, 112, 257, 628.

Ott, H. 1935 *Ann. Phys., Lpz.*, **23**, 169.

Raman, Sir C. V. 1941 *Proc. Indian Acad. Sci.* **13**, 1.

Raman, Sir C. V. & Krishnan, K. S. 1928 *Nature, Lond.*, **121**, 501, 771.

Raman, Sir C. V. & Nilakantan, P. 1941 *Curr. Sci.* **10**, 241.

Raman, Sir C. V. and collaborators 1941*a* Symposium of papers on the quantum theory of X-ray reflections. *Proc. Indian Acad. Sci.* **14**, 317.

Raman, Sir C. V. and collaborators 1941*b* Symposium of papers on the thermal energy of crystalline solids. *Proc. Indian Acad. Sci.* **14**, 459.

Robertson, Sir R., Fox, J. J. & Martin, A. E. 1934 *Phil. Trans.* A, **232**, 463.

Sarginson, K. 1941 *Proc. Roy. Soc.* A (in the Press), (quoted as Part II).

Smekal, A. 1923 *Naturwissenschaften*, **11**, 873.

Zachariasen, W. H. 1941 *Phys. Rev.* **57**, 597, 745; **59**, 860.

Thermodynamics of a relativistic Fermi-Dirac gas

By D. S. Kothari and B. N. Singh
University of Delhi

(*Communicated by M. N. Saha, F.R.S.—Received 5 August* 1941)

The thermodynamic properties—energy, pressure, free energies and entropy—for a Fermi-Dirac ideal gas are derived, taking account of the effect of relativistic mechanics. The results are expressed in convenient series-expansions. The degenerate and non-degenerate cases are considered.

1. In recent years the properties of a free-electron degenerate gas have been studied by Stoner and others in considerable detail, particularly when the effect of relativity mechanics is neglected. The relevant integrals for the completely non-relativistic case have been evaluated numerically by McDougall and Stoner (1938). There are, however, cases (e.g. in astrophysics) when the relativistic correction becomes important, and for this reason, and also for the sake of completeness, it appears worth while to derive general series-expansions for the important thermodynamic functions in terms of the non-dimensional parameter x,

$$x = \frac{h}{mc}\left(\frac{3n}{4\pi g}\right)^{\frac{1}{3}},$$

where n is the concentration and the other symbols have their usual meaning. The expressions given by Chandrasekhar (1939) do not involve x explicitly, and therefore are not so suitable for application as those which are obtained by formally extending to the general case Stoner's (1936) procedure of the non-relativistic case.*

The series-expansions in the degenerate case for the energy E, the pressure p, the Helmholtz free energy F, the Gibbs free energy G, and the constant-volume specific heat C_v are derived in the next section, taking

---

* The x occurring in Chandrasekhar's expressions is different from that defined here. Explicit series-expansions for thermodynamic functions, besides pressure and energy, have not been given by him. The connection between Chandrasekhar's results and ours can be seen in the following way:

Let y be the quantity called x by Chandrasekhar. Then formula (204), p. 392 of his book, can be written

$$x^3 = y^3\left[1 + \pi^2\left(\frac{kT}{mc^2}\right)^2 \frac{2y^2+1}{2y^4} + \dots\right],$$

where $x = \dfrac{h}{mc}\left(\dfrac{3}{8\pi}\dfrac{N}{V}\right)$ is our parameter (with $n = N/V$, $g = 2$). This equation must be solved for y and the result substituted in Chandrasekhar's formulae (198) and (205) for pressure and energy to give the expressions of this paper.

account of the effect of relativistic mechanics. The formulae for the two limiting cases, (i) the completely non-relativistic, and (ii) the completely relativistic, are given in table 1. The expressions for non-degeneracy in the two extreme cases* are also added at the end; these are obtained after obvious alterations in the corresponding formulae for the Bose-Einstein statistics given by Kothari and Singh (1941).

2. Let us consider an assembly of N similar particles occupying a volume V. Let ϵ represent the (translational) kinetic energy of a particle in the assembly, then the number of independent wave functions or states of a particle with eigen values of its kinetic energy lying in the range ϵ to $\epsilon + d\epsilon$ is, taking account of the effect of relativistic mechanics,

$$\frac{4\pi g V}{c^3 h^3}\, a(\epsilon)\, d\epsilon, \tag{1}$$

$$a(\epsilon) = (\epsilon^2 + 2\epsilon mc^2)^{\frac{1}{2}}\,(\epsilon + mc^2), \tag{2}$$

where c is the velocity of light, h Planck's constant, m the mass of a particle, and g its weight factor entering in virtue of its internal structure. Then, the Fermi-Dirac distribution law states that the number $N(\epsilon)\, d\epsilon$ of particles possessing energy between ϵ and $\epsilon + d\epsilon$ is given by

$$N(\epsilon)\, d\epsilon = \frac{4\pi g V}{c^3 h^3}\, \frac{a(\epsilon)\, d\epsilon}{e^{\epsilon/kT}/A + 1} = \frac{4\pi g V}{c^3 h^3}\, \frac{a(\epsilon)\, d\epsilon}{e^{(\epsilon - \zeta)/kT} + 1}, \tag{3}$$

where A or ζ ($\zeta = kT \log A$) is independent of ϵ, and as follows from the usual thermodynamical arguments, ζ is the Gibb's free energy per particle (chemical potential), and hence A may be called the *absolute activity*. For $\epsilon = \zeta$, $N(\epsilon)\, d\epsilon = \dfrac{4\pi g V}{c^3 h^3}\, \dfrac{a(\epsilon)\, d\epsilon}{2}$, that is, in the neighbourhood of ζ the number of energy states (of a particle) occupied by particles is half the available number of states.

If N denotes the total number of particles in the assembly, and E the total energy, then we have

$$N = \frac{4\pi g V}{c^3 h^3} \int_0^\infty \frac{(\epsilon^2 + 2mc^2 \epsilon)^{\frac{1}{2}}\,(\epsilon + mc^2)}{e^{(\epsilon - \zeta)/kT} + 1}\, d\epsilon, \tag{4}$$

$$E = \frac{4\pi g V}{c^3 h^3} \int_0^\infty \frac{\epsilon(\epsilon^2 + 2mc^2 \epsilon)^{\frac{1}{2}}\,(\epsilon + mc^2)}{e^{(\epsilon - \zeta)/kT} + 1}\, d\epsilon. \tag{5}$$

Let v represent the velocity and s the momentum of a particle, then by the usual arguments of the kinetic theory, the pressure p of the assembly will be given by

$$p = \frac{1}{3} \int_0^\infty \frac{vsN(\epsilon)\, d\epsilon}{V}, \tag{6}$$

---

* General expressions for non-degeneracy taking account of relativistic effect were first derived by Juttner (1928).

and substituting

$$v = \frac{c(\epsilon^2+2mc^2\epsilon)^{\frac{1}{2}}}{\epsilon+mc^2}, \quad \text{and} \quad s^2c^2 = \epsilon^2+2mc^2\epsilon,$$

we have
$$p = \frac{4\pi g}{3c^3h^3}\int_0^\infty \frac{(\epsilon^2+2mc^2\epsilon)^{\frac{3}{2}}}{e^{(\epsilon-\zeta)/kT}+1}\,d\epsilon. \tag{7}$$

An asymptotic series expansion for the integral

$$\int_0^\infty \frac{d\phi(\epsilon)}{d\epsilon}\,\frac{1}{e^{(\epsilon-\zeta)/kT}+1}\,d\epsilon,$$

when $\phi(\epsilon)$ is sufficiently regular and vanishes for $\epsilon = 0$, was first given by Sommerfeld. Subject to an error of the order of $e^{-\zeta/kT}$, an expansion appropriate for the case $\zeta/kT \gg 1$—the *degenerate case*—is

$$\int_0^\infty \frac{d\phi(\epsilon)}{d\epsilon}\,\frac{1}{e^{(\epsilon-\zeta)/kT}+1}\,d\epsilon$$
$$= \{\phi(\epsilon)+2c_2(kT)^2\phi''(\epsilon)+2c_4(kT)^4\phi^{\mathrm{iv}}(\epsilon)+\dots\}_{\epsilon=\zeta}, \tag{8}$$

where
$$c_{2n} = (1-2^{1-2n})\,\zeta(2n), \tag{9}$$

$\zeta(2n)$ being the Riemann-Zeta function. The numerical values of c_{2n} are (McDougall and Stoner 1938)

$$c_2 = \frac{\pi^2}{12} = 0{\cdot}822{,}467, \quad c_4 = \frac{7\pi^4}{720} = 0{\cdot}947{,}032, \quad c_6 = \frac{31\pi^6}{30240} = 0{\cdot}985{,}551.$$

If the assembly be considered to be reduced to the absolute zero of temperature (its volume being unchanged), then the distribution law (3) gives

$$N(\epsilon)\,d\epsilon = \frac{4\pi gV}{c^3h^3}\,a(\epsilon)\,d\epsilon \quad \text{for } \epsilon \leqslant \zeta_0,$$

$$N(\epsilon)\,d\epsilon = 0 \qquad\qquad \text{for } \epsilon > \zeta_0,$$

where ζ_0, the maximum energy of a particle in the completely degenerate case, is defined by

$$\frac{(ch)^3}{4\pi gV}N = \int_0^{\zeta_0} a(\epsilon)\,d\epsilon,$$
$$= \tfrac{1}{3}(\zeta_0^2+2mc^2\zeta_0)^{\frac{3}{2}}, \tag{10}$$

or
$$\zeta_0 = mc^2[(1+x^2)^{\frac{1}{2}}-1], \tag{11}$$

where
$$x = \frac{h}{mc}\left(\frac{3n}{4\pi g}\right)^{\frac{1}{3}}. \tag{12}$$

It may be noted for ready reference that

$$\zeta_0+mc^2 = mc^2(1+x^2)^{\frac{1}{2}} \quad \text{and} \quad (\zeta_0^2+2mc^2\zeta_0)^{\frac{1}{2}} = mc^2x.$$

The completely non-relativistic case is characterized by ζ_0 being negligible compared to the rest-mass energy, i.e. $x \to 0$; for the completely relativistic case $\zeta_0/mc^2 \to \infty$, i.e. $x \to \infty$.

We shall now proceed to obtain a series expansion for ζ. Using (8) in (4), and from (10) we have

$$\frac{(ch)^3}{4\pi gV}N = \int_0^{\zeta_0} a(\epsilon)\,d\epsilon,$$

$$= \int_0^{\zeta} a(\epsilon)\,d\epsilon + 2\left\{c_2(kT)^2\frac{da(\epsilon)}{d\epsilon} + c_4(kT)^4\frac{d^3a(\epsilon)}{d\epsilon^3} + \dots\right\}_{\epsilon=\zeta}, \tag{13}$$

and noting that

$$\int_0^{\zeta} a(\epsilon)\,d\epsilon = \int_0^{\zeta_0} a(\epsilon)\,d\epsilon + (\zeta-\zeta_0)\,a(\zeta_0) + \tfrac{1}{2}(\zeta-\zeta_0)^2\frac{da(\zeta_0)}{d\zeta_0} + \dots,$$

$$\left\{\frac{da(\epsilon)}{d\epsilon}\right\}_{\epsilon=\zeta} = \left\{\frac{da(\epsilon)}{d\epsilon}\right\}_{\epsilon=\zeta_0} + (\zeta-\zeta_0)\frac{d}{d\zeta_0}\left\{\frac{da(\epsilon)}{d\epsilon}\right\}_{\epsilon=\zeta_0} + \dots,$$

$$\left\{\frac{d^3a(\epsilon)}{d\epsilon^3}\right\}_{\epsilon=\zeta} = \left\{\frac{d^3a(\epsilon)}{d\epsilon^3}\right\}_{\epsilon=\zeta_0} + \dots,$$

we obtain after a little algebra, retaining terms up to the fourth power in temperature,

$$\zeta-\zeta_0 = -2c_2(kT)^2\frac{a'(\zeta_0)}{a(\zeta_0)} - 2(kT)^4\left\{c_4\frac{a'''(\zeta_0)}{a(\zeta_0)} + c_2^2\left(\frac{a'(\zeta_0)}{a(\zeta_0)}\right)^3 - 2c_2^2\frac{a'(\zeta_0)\,a''(\zeta_0)}{\{a(\zeta_0)\}^2}\right\}, \tag{14}$$

where

$$\left.\begin{aligned} a(\zeta_0) &= (mc^2)^2x(1+x^2)^{\frac{1}{2}},\\ a'(\zeta_0) &= mc^2\frac{1+2x^2}{x},\\ a''(\zeta_0) &= \frac{(1+x^2)^{\frac{1}{2}}(2x^2-1)}{x^3},\\ a'''(\zeta_0) &= \frac{1}{mc^2}\frac{3}{x^5}. \end{aligned}\right| \tag{15}$$

The distinction regarding differentiation with respect to ϵ and ζ_0 is ignored here and in all those cases where this leads to no ambiguity.* Substituting (15) in (14) we have

$$\zeta-\zeta_0 = -\frac{2c_2(kT)^2(1+2x^2)}{mc^2x^2(1+x^2)^{\frac{1}{2}}} - \frac{2(kT)^4}{(mc^2)^3x^6(1+x^2)^{\frac{1}{2}}}\left\{3c_4 + c_2^2\frac{(2x^2+1)(2x^2+3)}{1+x^2}\right\}, \tag{16}$$

or

$$\zeta = \zeta_0\left\{1 - \frac{\pi^2}{6}\left(\frac{kT}{\zeta_0}\right)^2\frac{(1+2x^2)[(1+x^2)^{\frac{1}{2}}-1]}{x^2(1+x^2)^{\frac{1}{2}}}\right.$$

$$\left. - \frac{\pi^4}{360}\left(\frac{kT}{\zeta_0}\right)^4\frac{(5x^2+4)(4x^2+9)[(1+x^2)^{\frac{1}{2}}-1]^3}{x^6(1+x^2)^{\frac{3}{2}}}\right\}. \tag{17}$$

---

* For entropy (equation (29)), the integrand contains a factor $(\epsilon-\zeta)$, and in that case the distinction mentioned above cannot be ignored.

We shall now determine the series-expansions for energy, pressure and entropy. Let us first treat the case of energy. From (1) and (4) we have

$$E = \frac{4\pi g V}{c^3 h^3} \int_0^\infty \frac{E(\epsilon)\, d\epsilon}{e^{(\epsilon-\zeta)/kT}+1}, \tag{18}$$

where $E(\epsilon) = \epsilon(\epsilon^2+2mc^2\epsilon)^{\frac{1}{2}}(\epsilon+mc^2) = \frac{1}{3}\epsilon \frac{d}{d\epsilon}(\epsilon^2+2mc^2\epsilon)^{\frac{3}{2}}$.

We obtain, therefore,

$$\begin{aligned}
\frac{E}{V}\frac{c^3h^3}{4\pi g} &= \int_0^\zeta E(\epsilon)\, d\epsilon + [2c_2(kT)^2 E'(\epsilon) + 2c_4(kT)^4 E'''(\epsilon)\ldots]_\zeta \\
&= \int_0^{\zeta_0} E(\epsilon)\, d\epsilon + (\zeta-\zeta_0)\, E(\zeta_0) + \frac{(\zeta-\zeta_0)^2}{2} E'(\zeta_0) \\
&\quad + [2c_2(kT)^2 E'(\zeta_0) + 2c_2(kT)^2(\zeta-\zeta_0)\, E''(\zeta_0) + 2c_4(kT)^4 E'''(\zeta_0)] + \ldots \\
&= \int_0^{\zeta_0} E(\epsilon)\, d\epsilon + [2c_2(kT)^2 E'(\zeta_0) + (\zeta-\zeta_0)\, E(\zeta_0)] + \Big[2c_4(kT)^4 E'''(\zeta_0) \\
&\quad + 2c_2(kT)^2(\zeta-\zeta_0)\, E''(\zeta_0) + \frac{(\zeta-\zeta_0)^2}{2} E'(\zeta_0)\Big] + \ldots,
\end{aligned} \tag{19}$$

where we have after integration and replacing ζ_0 by x,

$$\int_0^{\zeta_0} E(\epsilon)\, d\epsilon = (mc^2)^4 \left[\frac{x}{8}(1+x^2)^{\frac{1}{2}}(1+2x^2) - \frac{x^3}{3} - \frac{1}{8}\log\{x+(1+x^2)^{\frac{1}{2}}\}\right],$$

and

$$\left.\begin{aligned}
E(\zeta_0) &= (mc^2)^3 x[(1+x^2)-(1+x^2)^{\frac{1}{2}}], \\
E'(\zeta_0) &= \frac{(mc^2)^2}{x}[(1+x^2)^{\frac{1}{2}}(1+3x^2)-(1+2x^2)], \\
E''(\zeta_0) &= \frac{mc^2}{x^3}[6x^4+3x^2-1+(1-2x^2)(1+x^2)^{\frac{1}{2}}], \\
E'''(\zeta_0) &= \frac{3}{x^5}[(1+x^2)^{\frac{1}{2}}(2x^4-x^2+1)-1].
\end{aligned}\right\} \tag{20}$$

Substituting from (20) in (19) and rearranging in powers of kT, we have up to the fourth power in (kT)

$$\begin{aligned}
E = 3Nmc^2\Bigg[E(x) &+ 2c_2\left(\frac{kT}{mc^2}\right)^2 \frac{(1+x^2)^{\frac{1}{2}}}{x^2} \\
&+ 6\left(\frac{kT}{mc^2}\right)^4 \frac{(1+x^2)^{\frac{1}{2}}}{x^6}\left\{c_4(2x^2-1) - c_2^2\frac{(2x^2+1)^2}{1+x^2}\right\}\Bigg],
\end{aligned} \tag{21}$$

where $E(x)$ is defined by

$$E(x) = \frac{(1+x^2)^{\frac{1}{2}}(1+2x^2)}{8x^2} - \frac{1}{3} - \frac{1}{8x^3}\log\{(1+x^2)^{\frac{1}{2}}+x\}. \tag{22}$$

The asymptotic expansions for $E(x)$ are

$$\underset{x\to 0}{E(x)} = \frac{x^2}{10}-\frac{x^4}{56}+\frac{x^6}{144}-\frac{5x^8}{1408}+\ldots,$$

$$\underset{x\to\infty}{E(x)} = \frac{x}{4}-\frac{1}{3}+\frac{1}{4x}-\frac{1}{8x^3}\left(\log 2x-\frac{1}{4}\right)-\frac{1}{32x^5}+\ldots.$$

If C_v denotes the specific heat for constant volume, then we have

$$\begin{aligned}C_v &= \left(\frac{\partial E}{\partial T}\right)_v\\ &= 12Nk\left(\frac{kT}{mc^2}\right)\left[\frac{c_2(1+x^2)^{\frac{1}{2}}}{x^2}+6\left(\frac{kT}{mc^2}\right)^2\frac{(1+x^2)^{\frac{1}{2}}}{x^6}\left\{c_4(2x^2-1)-c_2^2\frac{(2x^2+1)^2}{1+x^2}\right\}\right].\end{aligned}$$

The completely relativistic and non-relativistic asymptotic expressions for E are at once written down from (21), and are given in table 1. We proceed now to derive the expression for the pressure of the assembly. Proceeding as in the case of the energy, we obtain from (7),

$$\begin{aligned}p\frac{3c^3h^3}{4\pi g} &= \int_0^\infty \frac{p(\epsilon)\,d\epsilon}{e^{(\epsilon-\zeta)/kT}+1}\\ &= \int_0^{\zeta_0} p(\epsilon)\,d\epsilon+[2c_2(kT)^2p'(\zeta_0)+(\zeta-\zeta_0)\,p(\zeta_0)]\\ &\qquad +\left[2c_4(kT)^4p'''(\zeta_0)+2c_2(kT)^2(\zeta-\zeta_0)\,p''(\zeta_0)\right.\\ &\qquad \left.+\frac{(\zeta-\zeta_0)^2}{2}p'(\zeta_0)\right]+\ldots,\end{aligned} \tag{23}$$

whence, after integrating and replacing ζ_0 in terms of x,

$$\left.\begin{aligned}\int_0^{\zeta_0} p(\epsilon)\,d\epsilon &= \int_0^{\zeta_0}(\epsilon^2+2mc^2\epsilon)^{\frac{3}{2}}\,d\epsilon\\ &= (mc^2)^4\tfrac{1}{8}[x(1+x^2)^{\frac{1}{2}}(2x^2-3)+3\log\{x+(1+x^2)^{\frac{1}{2}}\}],\\ p(\zeta_0) &= (mc^2x)^3,\\ p'(\zeta_0) &= (mc^2)^2\,3x(1+x^2)^{\frac{1}{2}},\\ p''(\zeta_0) &= (mc^2)\,3\frac{(1+2x^2)}{x},\\ p'''(\zeta_0) &= \frac{3(1+x^2)^{\frac{1}{2}}(2x^2-1)}{x^3}.\end{aligned}\right\} \tag{24}$$

Substituting from (24) in (23), and rearranging the terms, we have

$$\begin{aligned}p = \frac{N}{V}mc^2\Bigg[p(x)&+2c_2\left(\frac{kT}{mc^2}\right)^2\frac{2+x^2}{x^2(1+x^2)^{\frac{1}{2}}}\\ &+\frac{2}{x^6(1+x^2)^{\frac{1}{2}}}\left(\frac{kT}{mc^2}\right)^4\left\{3c_4(2x^4+x^2-2)-c_2^2\frac{(12x^6+28x^4+23x^2+6)}{1+x^2}\right\}\Bigg],\end{aligned} \tag{25}$$

where $p(x)$ is given by

$$p(x) = \frac{1}{8}\left[\frac{(1+x^2)^{\frac{1}{2}}(2x^2-3)}{x^2} + \frac{3}{x^3}\log\{(1+x^2)^{\frac{1}{2}}+x\}\right]; \tag{26}$$

and $$p(x)+3E(x) = (1+x^2)^{\frac{1}{2}}-1.$$

The asymptotic expressions are

$$\underset{x\to 0}{p(x)} = \frac{x^2}{5}-\frac{x^4}{14}+\frac{1}{24}x^6-\frac{5}{176}x^8+\dots$$

$$\underset{x\to\infty}{p(x)} = \frac{x}{4}-\frac{1}{4x}+\frac{3}{8x^3}\left(\log 2x-\frac{7}{12}\right)+\frac{5}{32x^5}-\dots.$$

We now take up entropy. If G denotes the Gibb's free energy and S the entropy, then

$$\left.\begin{aligned} G &= H-TS, \\ H &= E+pV. \end{aligned}\right\} \tag{27}$$

where

As already mentioned ζ represents the Gibb's free energy per particle, and therefore

$$G = N\zeta, \tag{28}$$

and substituting the integrals for N, E and p (equations (4), (5) and (7)), we have

$$TS = \frac{4\pi gV}{3c^3h^3}\int_0^\infty \frac{S(\epsilon)}{e^{(\epsilon-\zeta)/kT}+1}\,d\epsilon,$$

$$S(\epsilon) = \frac{d}{d\epsilon}\{(\epsilon-\zeta)(\epsilon^2+2mc^2\epsilon)^{\frac{3}{2}}\}. \tag{29}$$

The series-expansion for S can be immediately written down without actually evaluating the above integral, for we have already obtained expansions for E, p, and the expansion for G is from (16)

$$\begin{aligned} G = N\zeta = Nmc^2\bigg[\Big\{(1+x^2)^{\frac{1}{2}}-1\Big\} &- 2c_2\left(\frac{kT}{mc^2}\right)^2\frac{1+2x^2}{x^2(1+x^2)^{\frac{1}{2}}} \\ &- 2\left(\frac{kT}{mc^2}\right)^4\frac{1}{x^6(1+x^2)^{\frac{1}{2}}}\left\{3c_4+c_2^2\frac{(2x^2+1)(2x^2+3)}{(1+x^2)}\right\}\bigg], \end{aligned} \tag{30}$$

and substituting these in (27) we have

$$\begin{aligned} H = Nmc^2\bigg[\{(1+x^2)^{\frac{1}{2}}-1\} + 2c_2\left(\frac{kT}{mc^2}\right)^2\frac{5+4x^2}{x^2(1+x^2)^{\frac{1}{2}}} + 2\left(\frac{kT}{mc^2}\right)^4\frac{1}{x^6(1+x^2)^{\frac{1}{2}}} \\ \times\left\{3c_4(8x^4+4x^2-5)-c_2^2\frac{48x^6+100x^4+68x^2+15}{(1+x^2)}\right\}\bigg], \end{aligned} \tag{31}$$

and $$\begin{aligned} S = Nk\left(\frac{kT}{mc^2}\right)\bigg[\frac{\pi^2(1+x^2)^{\frac{1}{2}}}{x^2} + 24\left(\frac{kT}{mc^2}\right)^2\frac{1}{x^6(1+x^2)^{\frac{1}{2}}} \\ \times\{c_4(2x^2-1)(1+x^2)-c_2^2(2x^2+1)^2\}\bigg]. \end{aligned} \tag{32}$$

The Helmholtz free energy is defined by

$$F = G - pV, \tag{33}$$

and hence*

$$F = \frac{4\pi g V}{3c^3h^3}\int_0^\infty \frac{\zeta e^{\epsilon/\zeta}\dfrac{d}{d\epsilon}\left[e^{-\epsilon/\zeta}(\epsilon^2+2mc^2\epsilon)^{\frac{1}{2}}\right]}{e^{(\epsilon-\zeta)/kT}+1}\,d\epsilon, \tag{34}$$

and using the series expansions for G and p, we obtain

$$\begin{aligned} F = 3Nmc^2\Bigg[& E(x) - 2c_2\left(\frac{kT}{mc^2}\right)^2\frac{(1+x^2)^{\frac{1}{2}}}{x^2} \\ & - 2\left(\frac{kT}{mc^2}\right)^4 \frac{1}{x^6(1+x^2)^{\frac{1}{2}}}\{c_4(2x^2-1)(x^2+1)-c_2^2(2x^2+1)^2\}\Bigg]. \end{aligned} \tag{35}$$

We have obtained expressions taking account of relativistic mechanics for the different thermodynamic functions. It is usual to consider two extreme cases: (i) the *completely* non-relativistic case and (ii) the *completely* relativistic case. As already remarked, case (i) is characterized by $x \to 0$ and case (ii) by $x \to \infty$. Expressions for these limiting cases are easily derived from the general series-expansions already given, and the results are summarized in table 1. The expression for G in the relativistic case does not contain the fourth-power term in (kT), and for this limiting case (4) gives on integration, to an error of the order of $e^{-\zeta_0/kT}$,

$$\zeta^3 + 12c_2(kT)^2\zeta - \zeta_0^3 = 0,$$

or

$$\begin{aligned} \zeta &= 4(kT)\,c_2^{\frac{1}{2}}\sinh\left[\tfrac{1}{3}\sinh^{-1}\left\{\frac{1}{16c_2^{\frac{3}{2}}}\left(\frac{\zeta_0}{kT}\right)^3\right\}\right] \\ &= \zeta_0\left\{1 - 4c_2\left(\frac{kT}{\zeta_0}\right)^2 + \frac{(4c_2)^3}{3}\left(\frac{kT}{\zeta_0}\right)^6 + \dots\right]. \end{aligned}$$

For the sake of completeness the two limiting cases for non-degeneracy are also worth consideration. The formulae are easily obtained from the corresponding formulae (Kothari and Singh 1941) for the Bose-Einstein statistics after making some formal alterations. Thus we have

$$\begin{aligned} E &= sNkT[1 + b_2A_0 - b_3A_0^2 + b_4A_0^3 \dots], \\ pV &= NkT[1 + b_2A_0 - b_3A_0^2 + b_4A_0^3 \dots], \\ F &= NkT[-1 + \log A_0 + b_2A_0 - \tfrac{1}{2}b_3A_0^2 + \tfrac{1}{3}b_4A_0^3 \dots], \\ G &= NkT[\log A_0 + 2b_2A_0 - \tfrac{3}{2}b_3A_0^2 + \tfrac{4}{3}b_4A_0^3 \dots], \\ S &= Nk\left[-\log A_0 + (s+1) + (s-1)\,b_2A_0 - \frac{(2s-1)}{2}\,b_3A_0^3 + \frac{(3s-1)}{3}\,b_4A_0^3 \dots\right], \\ C_v &= sNk[1 - (s-1)\,b_2A_0 + (2s-1)\,b_3A_0^2 - (3s-1)\,b_4A_0^3 + \dots]. \end{aligned}$$

---

* It may be noted that $\left(\dfrac{\partial F}{\partial \zeta}\right)_{T,V} = 0.$

TABLE 1. DEGENERATE FERMI-DIRAC GAS $x = \frac{h}{mc}\left(\frac{3n}{4\pi g}\right)^{\frac{1}{3}}$

non-relativistic case	relativistic case
$\zeta_0 = \frac{h^2}{2m}\left(\frac{3n}{4\pi g}\right)^{\frac{2}{3}} = \frac{mc^2x^2}{2}$	$\zeta_0 = ch\left(\frac{3n}{4\pi g}\right)^{\frac{1}{3}} = mc^2x,$
$G = N\zeta = N\zeta_0\left[1 - c_2\left(\frac{kT}{\zeta_0}\right)^2 - \frac{3}{4}\alpha_1\left(\frac{kT}{\zeta_0}\right)^4\right]$	$G = N\zeta = N\zeta_0\left[1 - 4c_2\left(\frac{kT}{\zeta_0}\right)^2 + \frac{(4c_2)^3}{3}\left(\frac{kT}{\zeta_0}\right)^6\right]$
$E = \frac{3}{5}N\zeta_0\left[1 + 5c_2\left(\frac{kT}{\zeta_0}\right)^2 - \frac{15}{4}\alpha_1\left(\frac{kT}{\zeta_0}\right)^4\right]$	$E = \frac{3}{4}N\zeta_0\left[1 + 8c_2\left(\frac{kT}{\zeta_0}\right)^2 - 48\alpha_2\left(\frac{kT}{\zeta_0}\right)^4\right]$
$pV = \frac{2}{5}N\zeta_0\left[1 + 5c_2\left(\frac{kT}{\zeta_0}\right)^2 - \frac{15}{4}\alpha_1\left(\frac{kT}{\zeta_0}\right)^4\right]$	$pV = \frac{1}{4}N\zeta_0\left[1 + 8c_2\left(\frac{kT}{\zeta_0}\right)^2 - 48\alpha_2\left(\frac{kT}{\zeta_0}\right)^4\right]$
$H = N\zeta_0\left[1 + 5c_2\left(\frac{kT}{\zeta_0}\right)^2 - \frac{15}{4}\alpha_1\left(\frac{kT}{\zeta_0}\right)^4\right]$	$H = N\zeta_0\left[1 + 8c_2\left(\frac{kT}{\zeta_0}\right)^2 - 48\alpha_2\left(\frac{kT}{\zeta_0}\right)^4\right]$
$F = \frac{3}{5}N\zeta_0\left[1 - 5c_2\left(\frac{kT}{\zeta_0}\right)^2 + \frac{5}{4}\alpha_1\left(\frac{kT}{\zeta_0}\right)^4\right]$	$F = \frac{3}{4}N\zeta_0\left[1 - 8c_2\left(\frac{kT}{\zeta_0}\right)^2 + 16\alpha_2\left(\frac{kT}{\zeta_0}\right)^4\right]$
$S = 6Nk\left(\frac{kT}{\zeta_0}\right)\left[c_2 - \frac{1}{2}\alpha_1\left(\frac{kT}{\zeta_0}\right)^2\right]$	$S = 12Nk\left(\frac{kT}{\zeta_0}\right)\left[c_2 - 4\alpha_2\left(\frac{kT}{\zeta_0}\right)^2\right]$
$C_v = 6Nk\left(\frac{kT}{\zeta_0}\right)\left[c_2 - \frac{3}{2}\alpha_1\left(\frac{kT}{\zeta_0}\right)^2\right]$	$C_v = 12Nk\left(\frac{kT}{\zeta_0}\right)\left[c_2 - 12\alpha_2\left(\frac{kT}{\zeta_0}\right)^2\right]$

$$c_2 = \frac{\pi^2}{12} = 0{\cdot}822467, \qquad \left.\begin{aligned}\alpha_1 &= c_2^2 + c_4 = 1{\cdot}62348\\ \alpha_2 &= 2c_2^2 - c_4 = 0{\cdot}40587\end{aligned}\right\} \quad (\alpha_1 = 4\alpha_2).$$

For the non-relativistic case $s = \frac{3}{2}$, and

$$A_0 = \frac{nh^3}{g(2\pi mkT)^{\frac{3}{2}}} = \frac{4}{3\pi^{\frac{1}{2}}}\left(\frac{\zeta_0}{kT}\right)^{\frac{3}{2}}, \quad \zeta_0 = \frac{mc^2x^2}{2},$$

and for the relativistic case $s = 3$, and

$$A_0 = \frac{n}{8\pi g}\left(\frac{ch}{kT}\right)^3 = \frac{1}{6}\left(\frac{\zeta_0}{kT}\right)^3, \quad \zeta_0 = mc^2x.$$

The values of the coefficients b's are given below

$$b_2 = \frac{1}{2^{s+1}} = \left.\begin{matrix}0{\cdot}17678 & \text{for } s = \frac{3}{2},\\ 0{\cdot}06250 & \text{for } s = 3,\end{matrix}\right\}$$

$$b_3 = \frac{2}{3^{s+1}} - \frac{1}{4^s} = \left.\begin{matrix}0{\cdot}00330 & \text{for } s = \frac{3}{2},\\ 0{\cdot}009066 & \text{for } s = 3,\end{matrix}\right\}$$

$$b_4 = \frac{3}{4^{s+1}} + \frac{5}{2.8^s} - \frac{3}{6^s} = \left.\begin{matrix}0{\cdot}0001113 & \text{for } s = \frac{3}{2},\\ 0{\cdot}002713 & \text{for } s = 3.\end{matrix}\right\}$$

REFERENCES

Chandrasekhar, S. 1939 *An introduction to the study of stellar structure*, chapter 10.
Juttner, F. 1928 *Z. Phys.* **47**, 542.
Kothari, D. S. and Singh, B. N. 1941 *Proc. Roy. Soc.* A, **178**, 135.
Majumdar, R. C. 1931 *Astr. Nachr.* **242**, 145.
McDougall, J. and Stoner, E. C. 1938 *Phil. Trans.* A, **237**, 67.
Stoner, E. C. 1936 *Proc. Roy. Soc.* A, **154**, 656.

The photocombination of hydrogen and chlorine in oxygen-free systems

By M. Ritchie and D. Taylor

(*Communicated by J. Kendall, F.R.S.—Received 5 September* 1941)

The photosynthesis of hydrogen chloride has been studied under varied conditions of light intensity and of hydrogen, chlorine, hydrogen chloride, nitrogen and argon pressures. The inhibiting actions of chlorine and hydrogen chloride are considered confirmed. With increasing pressure of hydrogen chloride, the quantum efficiency decreased continuously, while with chlorine and also with nitrogen and argon, maxima were observed at intermediate pressures, after which the quantum efficiency fell progressively. With increasing light intensity at the lower pressures, the quantum efficiency was found to decrease to a constant value. Further, the change in intensity exponent with increasing pressure has been investigated, and values approaching 0·5 at high pressures have been observed. In the case of hydrogen chloride indices less than 0·5 were noted at the highest pressures.

These results are discussed on the basis of the following scheme:

$$
\begin{array}{ll}
Cl_2 \xrightarrow{h\nu} Cl + Cl, & Cl_3 + M \rightarrow Cl_2 + Cl + M, \\
Cl + H_2 \rightarrow HCl + H, & Cl_3 + Cl_3 \rightarrow 3Cl_2, \\
H + Cl_2 \rightarrow HCl + Cl, & Cl \rightarrow \text{surface}, \\
Cl + Cl_2 + M \rightarrow Cl_3 + M, & Cl_3 \rightarrow \text{surface}.
\end{array}
$$

To account for the abnormally low indices observed at high pressures, other chain ending processes are necessary; it is believed that the reactions

$$Cl + Cl_3 \rightarrow 2Cl_2 \quad \text{and} \quad Cl + Cl + M \rightarrow Cl_2 + M$$

are of importance in this connexion.

While the main outlines of the kinetics of the hydrogen-chlorine photocombination in oxygen-rich mixtures appear to be well established, it cannot be said that a comprehensive reaction scheme has been evolved which will definitely correlate the results of the many investigations on the corresponding oxygen-free systems. This communication presents the results of some further work in this latter field.

There seems little doubt as to the nature of the initiation and propagation of the chains by which the reaction proceeds. For light of the blue-violet region of the spectrum, as used in this investigation, the absorption of a quantum of radiation by a chlorine molecule results in the direct production of two chlorine atoms, and the actual chain is represented by the Nernst mechanism

$$\mathrm{Cl} + \mathrm{H_2} \rightarrow \mathrm{HCl} + \mathrm{H}, \quad \mathrm{H} + \mathrm{Cl_2} \rightarrow \mathrm{HCl} + \mathrm{Cl}.$$

The problem is then the determination of the processes which bring these chains to an end. Normally, the mechanism of the ultimate removal of chlorine atoms is of prime importance.

The possibility of Cl_3 molecules as intermediates in this connexion was considered by Göhring (1921) and by Cremer (1927), but little further use was made of it in kinetic schemes until Rollefson & Eyring (1932) showed on quantum-mechanical grounds that the Cl_3 quasi-molecule, formed from a chlorine atom and a chlorine molecule, should have a very long life as compared with the duration of an ordinary collision. Craggs, Squire & Allmand (1937) interpreted the results of an extensive series of experiments on this basis; on the other hand, Bodenstein, Brenschede & Schumacher (1938) have remarked that there are no grounds for the assumption of Cl_3 as an intermediate of importance in the hydrogen chloride photosynthesis, from the study of the other numerous investigations involving illuminated chlorine, and that the results of Allmand and co-workers are to be explained on other lines.

Ritchie & R. L. Smith (1940) concluded from an examination of the photo-expansion of chlorine, alone and in the presence of different pressures of other added gases, that in the gas phase chlorine atoms are removed by the triple collision process

$$\mathrm{Cl} + \mathrm{Cl_2} + M \rightarrow \mathrm{Cl_3} + M,$$

where M represents the stabilizing added gas molecule. If Cl_3 molecules are in turn removed by mutual recombination ($\mathrm{Cl_3} + \mathrm{Cl_3} \rightarrow 3\mathrm{Cl_2}$), the presence of such gases and of the product hydrogen chloride would mean under certain conditions a smaller quantum efficiency for the overall hydrogen chloride photosynthesis. Ritchie & Norrish (1933) recorded a marked decrease as the hydrogen chloride concentration was increased, but considered as responsible the reaction

$$\mathrm{H} + \mathrm{HCl} \rightarrow \mathrm{H_2} + \mathrm{Cl}.$$

It appears, however, that this reaction cannot have the efficiency necessary to account for the observed retardation (see Potts & Rollefson 1935). According to Bodenstein & Winter (1936) all possible effects (acceleration, no effect, retardation) could be obtained with different hydrogen chloride pressures according to the catalytic activity of the reaction vessel walls. Similarly, Craggs *et al.* (1937) attributed a rise in quantum efficiency as the chlorine pressure increased over a certain range to decreasing wall activity; this explanation was also applied to accelerations found with increasing concentrations of hydrogen chloride, while the suggestion was made that retardation was due to increased formation of the more easily adsorbed Cl_3 molecules. The elucidation of the kinetics of this complicated reaction must therefore involve discussion of the relative rates of diffusion of possible intermediates through the various gas mixtures, as well as of the nature of the surfaces removing these intermediates by adsorption; in addition, for a reaction such as this, in which a considerable amount of heat is evolved, possible effects of convection must be included. It is believed that the results here recorded do allow some distinction between these possibilities; they are, in fact, interpreted by the Cl_3 hypothesis.

If the main chain ending processes are represented by the formation of Cl_3 by a triple collision as above, and the removal of Cl_3 molecules by mutual recombination, then certain conditions may be expected to arise in which the rate of reaction on this account alone should become directly proportional to the intensity of absorbed light ($I_{\text{abs.}}$). If Cl_3 molecules decompose reversibly, the competition between these processes is such that at low intensities the rate may be proportional to $I^n_{\text{abs.}}$, where n approaches 0·5, but at high intensities n should approach unity. Preliminary experiments, by R. L. Smith in this Department, did in fact show a trend in this direction. Also, while the rate was much less at higher pressures of hydrogen chloride, the index n, in the neighbourhood of unity at low hydrogen chloride pressures and high values of $I_{\text{abs.}}$, became steadily less as the concentration of hydrogen chloride was increased. Accordingly these points have been systematically examined; the effect of an extensive variation of light intensity has been recorded, as well as the effects of altered concentrations of chlorine, hydrogen, hydrogen chloride, nitrogen and argon on the overall rate of reaction.

Experimental procedure and results

The experimental arrangement was similar in outline to that described by Ritchie & Norrish (1933). Light from a mercury vapour Osira lamp (125 W, 220 V by Solus stabilizer), rendered parallel by means of a suitable system of lenses and circular apertures, was directed in such a way as to fill almost completely the particular reaction vessel in use. Light of approximate wave-lengths 3650 and 4060 A was employed; for the 3650 A region, the filter was a 3 mm. plate of Chance's 'ultra-violet' glass, while for 4060 A, a 2 cm. layer of 0·1 % quinine hydrochloride aqueous solution was used in conjunction with a 2 cm. layer of 0·03 N iodine solution in carbon tetrachloride. In both cases the light from the lamp passed through a 3 cm. layer of 6 % aqueous copper sulphate solution before reaching the filters. Behind the reaction vessel and covering the entire area of the light beam was a Weston Photronic cell; this with a mirror galvanometer and scale recorded the light intensity as galvanometer scale divisions. Pressures in the reaction vessel were measured by a balanced Bourdon gauge; both vessel and gauge were water-jacketed to 25 ± 0·1° by a circulating water system. Control experiments with conditions similar to that of any actual insolation apart from actual insolation itself showed no change in chlorine concentration over periods of time much greater than that involved in actual rate determinations. A Hyvac oil pump, protected by liquid oxygen traps, was used in all operations.

Great care was taken in the purification of the reactant gases. Cylinder chlorine was passed through alkaline permanganate, concentrated sulphuric acid, and phosphoric oxide before being condensed by liquid oxygen and purified by repeated distillation, with rejection of initial and end fractions. When not in use, the product was kept solid by liquid oxygen. Cylinder argon and nitrogen were passed over copper at 400°, phosphoric oxide, through molten sodium and through liquid oxygen traps. Hydrogen was prepared electrolytically and purified by heated copper and palladiumized asbestos, followed by passage over a white-hot tungsten filament, with suitable phosphoric oxide and liquid oxygen traps. Hydrogen chloride was always prepared in the reaction vessel itself, by illumination of suitable hydrogen-chlorine mixtures.

Two cylindrical reaction vessels were employed, of clear silica, with polished plane end-plates; the diameter of each was approximately equal to the length. The rates of reaction were determined by the method of averaging described by Ritchie & Norrish, purified cylinder oxygen being added at the end of the time of insolation to render the mixture insensitive when the resultant chlorine concentration was determined. Absolute values of quantum efficiency, and the quanta absorbed for each wave-length for the particular vessel in use, were estimated by comparison with Norrish & Ritchie's absolute values for corresponding oxygen-rich mixtures (1933), on the assumption that the efficiency depends only on the concentrations of reactants (see Discussion).

Certain precautions in the experimental procedure must be mentioned. On evacuation of the system by oil pump care was taken that as far as possible no air previously in contact with the mercury of the direct-reading manometer was allowed into the reaction vessel. During the operations of washing out and filling the vessel with chlorine to any desired pressure, illumination was applied to destroy any chance inhibitors. A small correction was applied in the case of the small reaction vessel by reason of the chlorine pushed out from the connecting capillary tubing by the subsequently entering gases; no such correction was necessary for the large vessel. The values of the incident light intensity tended to decrease slightly during an experiment; any small change which did occur was taken as linear with respect to the time. The apiezon grease L used on all taps was satisfactory, all unnecessary greasing being avoided. The rate of the first experiment after the regreasing of the tap through which the gases were admitted to the reaction vessel was sometimes slightly lower than the standard value for the given conditions. With daily experiment regreasing was usually necessary at three-weekly intervals. Draper effects were normally sharp and immediate and could be taken as a reliable indication of the absence of abnormalities.

When the system was first set up, it was found that concordant results could not be obtained, but constantly increasing rates were recorded. These increasing values, however, reached a maximum which could be reproduced with a normal experimental error of less than 5 %, provided the system was not disturbed further. At the end of several series of experiments, the previously transparent silica surfaces were observed to show a slight opalescence, which could not be removed by treatment with concentrated nitric acid. On removal by shaking with clean sand, the rate of reaction fell to a considerably lower figure. Repeated experiment, however, restored the rate to its former value, no opacity being visible in the remaining experiments; surface factors were obviously in operation.

Experimental results are given in the following tables. Quantum efficiencies were calculated by the formula

$$\gamma_{\text{HCl}} = \frac{k'2[\Delta \text{Cl}_2]}{t I_{\text{abs.}}},$$

where ΔCl_2 is the change in chlorine concentration produced by absorption of $I_{\text{abs.}}$ quanta during t sec., and k' is a constant, depending on the reaction vessel and the wave-length used, and chosen to bring the oxygen-rich efficiencies to concordance with the corresponding absolute efficiencies determined by Norrish & Ritchie. Typical calibration series are given in tables 1, 2 and 3: small but apparently definite variations in k' occurred from time to time, but corresponding values were used in the calculations for oxygen-free mixtures. The absolute values of γ_{HCl} according to Norrish & Ritchie are 124 and 560 for the corresponding conditions.

On such a calibration basis, the efficiencies for oxygen-free mixtures are given in tables 3–10, where $I_{\text{abs.}}$ is now given in total quanta per second absorbed throughout the reaction vessel. In table 3, γ'_{HCl} is the observed quantum efficiency and γ_{HCl} is the efficiency calculated on the assumption of direct proportionality to hydrogen concentration. Since the concentrations of chlorine and hydrogen chloride and the value of $I_{\text{abs.}}$ remain essentially constant throughout, the formula $\gamma_{\text{HCl}} = C[\text{H}_2]$, where $C = 3{\cdot}24 \times 10^2$, has been applied. The departure from linearity as shown by the ratios of the last column is negligible at pressures up to 300 mm. hydrogen, and it is concluded that, in conformity with the results of other workers, the quantum efficiency is directly proportional to the pressure of hydrogen for these lower pressures. In tables 4–10 the small variations in $[H_2]$ have

TABLE 1. SMALL REACTION VESSEL

Volume = 31·3 c.c.

date	$[O_2]$ mm.	$[H_2]$ mm.	$[Cl_2]$ mm.	$[HCl] = [\Delta Cl_2]$ mm.	$I_{abs.}$ scale div.	t sec.	γ_{HCl}
			$\lambda = 3650$ A; $k' = 9{\cdot}47 \times 10^4$				
14. ii. 40	50·3	43·45	44·45	13·5	13·93	1440	127·5
	50·5	42·8	42·4	15·2	13·44	1800	119·5
	50·8	42·5	41·7	15·4	13·01	1800	125·0
	50·7	43·1	42·4	15·2	12·96	1800	123·5
			$\lambda = 3650$ A; $k' = 8{\cdot}99 \times 10^4$				
25. v. 40	51·2	42·75	42·75	16·3	12·91	1800	126·0
	50·7	42·85	43·25	15·1	12·88	1680	125·5
	51·2	43·85	42·65	15·1	13·0	1680	124·5
	51·1	43·35	42·55	15·7	13·0	1800	121·0
			$\lambda = 4060$ A; $k' = 14{\cdot}7 \times 10^4$				
2. iii. 40	10·0	42·95	42·75	15·9	1·73	4800	562
	10·0	42·85	42·85	15·3	1·68	4800	556
			$\lambda = 4060$ A; $k' = 14{\cdot}52 \times 10^4$				
25. v. 40	10·0	42·6	41·9	16·4	2·40	3540	560

TABLE 2. LARGE REACTION VESSEL

Volume = 178·0 c.c.; $\lambda = 4060$ A; $k' = 17{\cdot}5 \times 10^4$

date	$[O_2]$ mm.	$[H_2]$ mm.	$[Cl_2]$ mm.	$[HCl] = [\Delta Cl_2]$ mm.	$I_{abs.}$ scale div.	t sec.	γ_{HCl}
4. vi. 40	10·0	44·95	44·45	12·5	3·7	2130	555
	10·0	44·45	44·15	13·1	3·61	2280	538
	10·0	43·9	43·8	14·4	3·48	2580	562

TABLE 3. VARIATION OF HYDROGEN PRESSURE

Vessel volume = 31·3 c.c.; $\lambda = 4060$ A; $k' = 14{\cdot}7 \times 10^4$

$[H_2]$ mm.	$[Cl_2]$ mm.	$[HCl] = [\Delta Cl_2]$ mm.	t sec.	$I_{abs.} \times 10^{-13}$ quanta/sec.	$\gamma'_{HCl} \times 10^{-4}$	$\gamma_{HCl} \times 10^{-4}$	$\gamma_{HCl}/\gamma'_{HCl}$
42·3	42·8	16·4	220	11·17	1·41	1·37	0·97
93·0	42·4	15·8	100	11·6	2·77	3·02	1·09
189·0	41·45	17·3	50	11·3	6·23	6·12	0·98
279·2	42·4	15·2	30	11·45	9·02	9·05	1·00
395·4	42·7	16·4	27	11·3	10·90	12·80	1·18
506·3	42·55	16·3	23	11·6	12·50	16·40	1·31
570·2	41·35	18·7	24	11·5	13·80	18·50	1·34

therefore been corrected to $[H_2] = 43$ mm. on this basis and are given in the columns under the heading γ_{HCl}. If it is assumed that the 'non-varying' parameters can be taken as constant, the effect of variation of $I_{abs.}$ for the

TABLE 4. VARIATION OF $I_{abs.}$

Vessel volume = 31·3 c.c.; $\lambda = 4060$ A; $k' = 14{\cdot}7 \times 10^4$;
$\lambda = 3650$ A; $k' = 9{\cdot}47 \times 10^4$

A	$[H_2]$ mm.	$[Cl_2]$ mm.	[HCl] mm.	t sec.	$I_{abs.} \times 10^{-12}$ quanta/sec.	$\gamma'_{HCl} \times 10^{-4}$	$\gamma_{HCl} \times 10^{-4}$
3650	43·0	42·4	15·2	17	153·4	1·18	1·18
3650	43·0	42·5	15·75	18	149·0	1·20	1·20
3650	43·7	43·7	14·6	32	78·5	1·18	1·16
3650	43·1	41·4	16·6	110	25·4	1·21	1·21
3650	43·2	43·4	13·6	100	21·4	1·28	1·28
3650	44·6	43·2	14·2	120	17·85	1·35	1·30
4060	43·0	42·8	15·3	155	15·8	1·28	1·28
4060	42·8	42·7	15·4	210	10·85	1·38	1·39
3650	43·4	42·7	15·4	270	8·10	1·44	1·42
4060	42·7	42·0	17·2	420	5·29	1·58	1·60
3650	41·45	41·85	17·9	540	4·36	1·54	1·60
4060	43·05	42·05	16·3	590	3·25	1·73	1·73
3650	42·9	42·2	16·2	840	2·20	1·78	1·785
4060	43·45	42·95	14·9	990	1·47	2·08	2·06
4060	42·9	41·6	16·6	1800	0·855	2·20	2·20
4060	43·05	42·05	15·9	3900	0·306	2·73	2·73

TABLE 5. VARIATION OF $I_{abs.}$

$[H_2]$ mm.	$[Cl_2]$ mm.	[HCl] mm.	t sec.	$I_{abs.} \times 10^{-12}$ quanta/sec.	$\gamma'_{HCl} \times 10^{-4}$	$\gamma_{HCl} \times 10^{-4}$
Vessel volume = 178·0 c.c.; $\lambda = 4060$ A; $k' = 17{\cdot}5 \times 10^4$						
43·0	42·95	15·5	45	118	3·34	3·34
42·0	42·6	17·0	50	117	3·34	3·42
43·1	42·95	15·7	45	112·8	3·56	3·55
43·5	42·95	15·1	62	81·7	3·43	3·40
43·3	43·2	15·2	62	79	3·57	3·54
43·0	42·6	15·8	155	34·1	3·44	3·44
43·7	43·7	14·2	280	16·23	3·60	3·54
43·1	42·7	16·2	520	8·99	4·00	3·99
43·1	42·85	15·7	780	4·98	4·66	4·65

TABLE 6. VARIATION OF CHLORINE PRESSURE

Vessel volume = 31·3 c.c.; $\lambda = 4060$ A; $k' = 14{\cdot}5 \times 10^4$

43·0	15·2	7·4	270	5·89	0·95	0·95
43·4	30·7	10·8	165	11·65	1·15	1·14
42·3	42·6	16·2	155	16·3	1·30	1·33
42·9	67·4	27·0	180	24·8	1·24	1·24
46·15	86·5	27·1	135	32·7	1·25	1·17
43·2	93·3	16·2	78	35·0	1·21	1·20
43·8	102·0	38·2	180	35·8	1·20 (5)	1·18

small vessel (table 4) and the effects of addition of argon and hydrogen chloride (tables 8, 9) are shown in figures 1 and 2 respectively. For clarity the nitrogen curves have been omitted; they occupy positions intermediate between those of argon and hydrogen chloride.

TABLE 7. VARIATION OF CHLORINE PRESSURE

$[H_2]$ mm.	$[Cl_2]$ mm.	[HCl] mm.	t sec.	$I_{abs.} \times 10^{-12}$ quanta/sec.	γ'_{HCl} $\times 10^{-4}$	γ_{HCl} $\times 10^{-4}$
Vessel volume = 178·0 c.c.; $\lambda = 4060$ A; $k' = 17{\cdot}5 \times 10^4$						
42·35	15·75	8·7	75	46·2	2·88	2·94
43·15	30·25	10·1	40	85·9	3·38	3·36
42·4	42·4	16·2	45	117·3	3·53	3·58
43·2	67·2	27·0	45	174·5	3·95	3·94
44·8	103·7	34·1	130	85·6	3·53	3·38
66·9	204·6	71·1	200	147	2·78	1·79
107·6	308·9	91·7	420	83·5	3·01	1·21

TABLE 8. VARIATION OF HYDROGEN CHLORIDE PRESSURE

Vessel volume = 31·3 c.c.

(a) $\lambda = 3650$ A; $k' = 8{\cdot}99 \times 10^4$						
43·1	43·1	15·0	17	149	1·21	1·21
42·85	42·35	71·5	30	153	0·728	0·731
44·0	43·0	122·4	38	155	0·510	0·497
42·3	41·6	123·8	45	148	0·540	0·550
43·4	43·3	226·4	80	155	0·240	0·240
43·6	43·5	313·2	135	153	0·150	0·147
44·75	45·05	410·3	150	162	0·099	0·095
44·7	44·9	434·3	120	152	0·141	0·136
(b) $\lambda = 4060$ A; $k' = 14{\cdot}5 \times 10^4$						
41·0	41·3	16·8	155	16·6	1·33	1·39
42·75	42·35	66·3	180	17·2	1·11	1·12
44·25	42·85	115·3	180	17·1	0·93	0·905
43·55	43·25	216·3	225	17·4	0·756	0·748
44·35	44·05	266·9	240	17·8	0·644	0·625
40·75	39·75	358	500	15·8	0·546	0·578
40·45	38·95	358	500	15·7	0·552	0·587
43·7	41·8	421	345	17·1	0·527	0·519

DISCUSSION

Mention has already been made of the fact that results were not at all times perfectly reproducible, more especially when the apparatus had been dismantled and the reaction vessel internally cleaned. This has been attributed to surface factors involving the removal of chain carriers, perhaps the best example of such 'disturbance' being the greatly decreased quantum efficiency observed when the slight wall opalescence, apparent after a

TABLE 9. VARIATION OF ARGON PRESSURE

Vessel volume = 31·3 c.c.

[A]	$[H_2]$ mm.	$[Cl_2]$ mm.	[HCl] mm.	t sec.	$I_{abs.} \times 10^{-12}$ quanta/sec.	$\gamma'_{HCl} \times 10^{-4}$	$\gamma_{HCl} \times 10^{-4}$
		(*a*) $\lambda = 3650$ A;	$k' = 8{\cdot}99 \times 10^4$				
0	43·1	43·1	15·0	17	147	1·21	1·21
53·8	42·65	42·35	16·7	18	149	1·28	1·285
154·9	43·1	42·5	16·4	18	147	1·27	1·27
207·4	42·15	41·55	18·5	21	149	1·21	1·24
306·7	42·65	42·35	16·7	23	148	1·01	1·02
407·2	43·1	42·7	16·0	28	149	0·79	0·79
		(*b*) $\lambda = 4060$ A;	$k' = 14{\cdot}5 \times 10^4$				
0	42·3	42·6	16·2	155	16·4	1·30	1·33
73·2	43·65	43·15	15·5	120	16·8	1·57	1·54
152·7	43·3	42·9	15·8	120	16·5	1·63	1·61
254·1	43·6	43·1	15·2	120	16·8	1·54	1·53
354·2	43·25	42·35	16·3	150	16·8	1·32	1·31
398·7	41·3	41·7	18·6	190	16·7	1·19	1·24

TABLE 10. VARIATION OF NITROGEN PRESSURE

Vessel volume = 31·3 c.c.

$[N_2]$							
		(*a*) $\lambda = 3650$ A;	$k' = 8{\cdot}99 \times 10^4$				
0	43·3	43·2	15·2	22	137	1·03	1·02
0	43·85	42·95	15·7	22	138	1·06	1·04
0	42·5	42·9	15·6	21	141	1·08	1·10
54·3	43·6	42·6	14·8	21	139	1·04	1·02
105·8	43·3	41·5	16·4	23	137	1·06	1·05
203·7	42·95	42·55	16·3	28	143	0·834	0·835
204·4	43·1	43·0	15·6	28	140	0·815	0·815
304·3	42·15	40·95	17·1	40	136	0·643	0·657
407·9	42·95	40·25	15·3	48	132	0·496	0·496
		(*b*) $\lambda = 4060$ A;	$k' = 14{\cdot}5 \times 10^4$				
0·0	43·4	42·8	16·0	160	15·9	1·29	1·28
0·0	43·05	43·05	15·7	155	15·9	1·30	1·30
0·0	42·95	43·15	15·1	155	15·8	1·26	1·26
73·7	43·15	42·55	16·7	150	15·7	1·45	1·44
154·0	43·65	43·15	14·7	120	16·1	1·56	1·53
255·7	43·35	42·35	15·5	150	15·8	1·335	1·33
351·7	42·95	41·05	15·9	210	15·0	1·03	1·03
404·1	42·35	41·05	16·7	225	16·7	0·905	0·916

considerable number of experiments, was removed. Such opalescence under similar conditions has been previously noted by other workers: Bodenstein & Unger (1931) cite experiments by Stieger and by Hertel as showing that on sufficiently prolonged illumination, chlorine atoms attack quartz, with the production of a substance which yields hydrochloric and silicic acids on treatment with water. We therefore attribute the faint opalescence here observed to such surface action rather than to any impurity introduced with the reactants. In general, constant results for given conditions of concentration and light intensity were obtained after repeated experiments, and, once obtained, could be reproduced with an error of less than 5 %, provided no further disturbance of the apparatus was involved. The quantum efficiencies recorded in the preceding tables all refer to such stationary conditions, and are therefore directly comparable, with perhaps one exception—the γ-N_2 curve at high $I_{\text{abs.}}$ (see table 10*a*); the initial value for $[N_2] = 0$ is some 12 % less than the corresponding points in the argon and hydrogen chloride series. The results in this nitrogen series were at the time reproducible within the limits of the other experimental errors.

Study of the variation of $I_{\text{abs.}}$ has been accomplished by making use of light of two wave-lengths, such a procedure being necessary because of the absorption coefficients concerned and the method of calculating the rates of reaction. For this purpose it has been assumed that the quantum efficiency is here independent of the wave-length; for the small change involved (3650–4060 A), it is unlikely that any appreciable error results from such an assumption. Ritchie & Norrish (1932) found in fact that efficiencies, obtained by a technique similar to that used here, were only 5 % less at 4060 A than those at 3650 A; this difference may be entirely due to experimental error. Accordingly, it is believed that the present results for the 500-fold variation of $I_{\text{abs.}}$ here investigated may be justifiably considered on a common basis.

As already explained, the absolute magnitudes of the quantum efficiencies have been estimated by comparison for oxygen-rich mixtures with the results of Norrish & Ritchie. Such a comparison can be justified only if surface action is absent or is the same for both systems, and if the efficiency is independent of the amount of light absorbed. Actually the same general technique is concerned, with transparent silica vessels subjected to similar treatment; while some objection is possible by reason of the fact that somewhat larger vessels were here used, a similar procedure applied to quantum efficiencies of water formation (Ritchie 1937) in a glass vessel gave results in good agreement with theoretical values, this indicating that in such oxygen-rich mixtures surface effects are inappreciable. Similarly in oxygen-rich mixtures it is an accepted fact that the efficiency is independent of the amount of light absorbed. While, therefore, relative efficiencies are in general sufficient for the following analysis of the present results, it is believed that the absolute values are not far from the magnitudes quoted.

The main trends which are indicated by the results of the previous tables, and which must be accounted for by any scheme of reaction mechanism are then

(*a*) The decrease in quantum efficiency to a constant value, independent of $I_{\text{abs.}}$, as $I_{\text{abs.}}$ is increased (figure 1, tables 4, 5).

(*b*) The increase in quantum efficiency as the chlorine concentration is increased, this being succeeded by a fall in efficiency at higher chlorine pressures (tables 6, 7).

(*c*) The initial increases in quantum efficiency as argon and nitrogen are added (figure 2, tables 9, 10).

(*d*) The decreases in quantum efficiency at higher pressures of argon and nitrogen, also observed with increasing pressure of hydrogen chloride (figure 2, tables 9, 10, 8).

Further, if the relation between the rate of reaction and the amount of light absorbed, other factors remaining constant, is given by

$$\frac{d(\mathrm{HCl})}{dt} = kI^{n}_{\mathrm{abs.}},$$

comparison of efficiencies $\left(\gamma_{\mathrm{HCl}} = \frac{1}{I_{\mathrm{abs.}}}\frac{d(\mathrm{HCl})}{dt}\right)$ at corresponding pressures of A, N_2, and HCl leads to the following calculated values of n (table 11). Quantum efficiencies have been taken from the smoothed curves (cf. figure 2): $I_{\mathrm{abs.}}$ represents average values throughout the series.

TABLE 11

Argon mm.	γ_{HCl} $I_{\mathrm{abs.}} = 149 \times 10^{12}$	γ_{HCl} $I_{\mathrm{abs.}} = 16{\cdot}7 \times 10^{12}$	n
0	12,100	13,300	0·96
100	12,900	16,000	0·90
200	12,200	16,000	0·88
300	10,300	14,400	0·85
400	8,200	12,600	0·80
nitrogen mm.	γ_{HCl} $I_{\mathrm{abs.}} = 139 \times 10^{12}$	γ_{HCl} $I_{\mathrm{abs.}} = 15{\cdot}9 \times 10^{12}$	n
0	10,500	12,900	0·90
100	10,100	14,700	0·83
200	8,400	14,200	0·76
300	6,600	11,600	0·74
400	4,800	9,600	0·68
Hydrogen chloride mm.	γ_{HCl} $I_{\mathrm{abs.}} = 153 \times 10^{12}$	γ_{HCl} $I_{\mathrm{abs.}} = 16{\cdot}8 \times 10^{12}$	n
15	12,100	13,800	0·94
100	6,300	9,700	0·80
200	3,000	7,300	0·60
300	1,600	6,000	0·40
400	1,400	5,400	0·39

To the trends (*a*), (*b*), (*c*) and (*d*) above, we therefore add (*e*), the change in index n to smaller values as argon, nitrogen and hydrogen chloride are added.

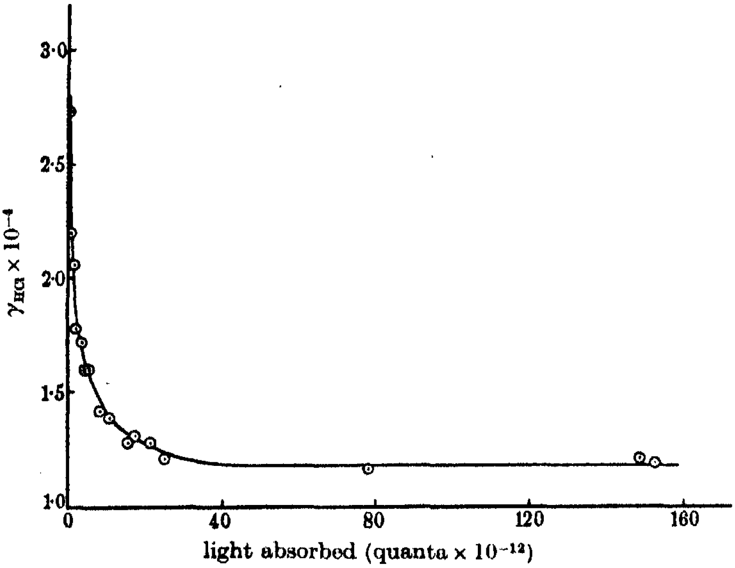


FIGURE 1

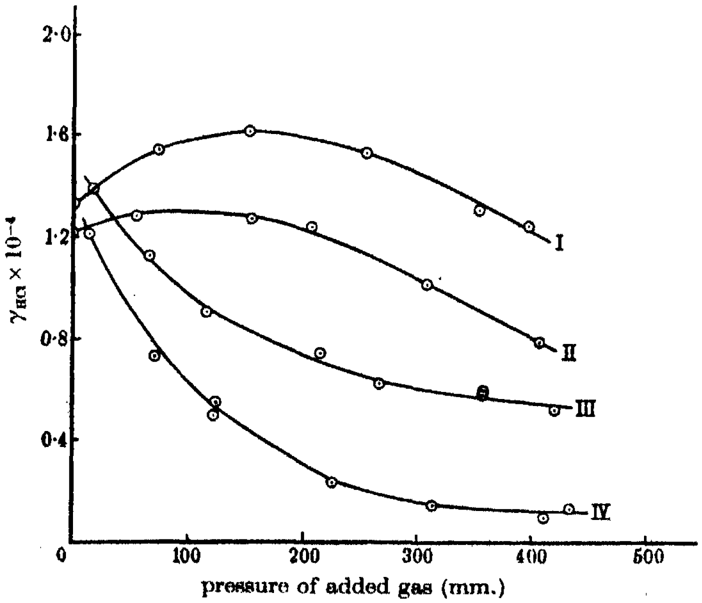


FIGURE 2

curve I, argon, $I_{abs.} = 16{\cdot}7 \times 10^{12}$; curve II, argon, $I_{abs.} = 149 \times 10^{12}$; curve III, hydrogen chloride, $I_{abs.} = 16{\cdot}8 \times 10^{12}$; curve IV, hydrogen chloride, $I_{abs.} = 153 \times 10^{12}$.

It would appear that apart from the Cl_3 mechanism later discussed in detail, there are two main ways in which such results may be considered. First, the combination of hydrogen and chlorine is attended by the evolution of a considerable quantity of heat. Thus, if the rate of reaction be large, as it is in some of these experiments, convection effects may play an important part. It is, for example, possible that an increased $I_{\text{abs.}}$ will cause considerably increased convection by reason of the greater heat evolution, and such an increased convection will tend to increase surface removal of chain carriers with a consequent decrease in quantum efficiency in the sense shown by figure 1. The change in index here, from values much less than 1 to a value of unity at high $I_{\text{abs.}}$, would on this basis be due to the progressive removal of all chain carriers by the walls, the quantum efficiency being finally independent of $I_{\text{abs.}}$. Evidence for a similar mechanism has been given for the reaction between illuminated bromine and acetylene (Müller & Schumacher 1938). The decrease at high chlorine concentrations might be accounted for in a similar way. On the other hand, increase in the hydrogen chloride concentration for these conditions of high $I_{\text{abs.}}$ causes a marked decrease in efficiency (figure 2), for conditions where all chain carriers are already presumably removed by surface action, such addition of hydrogen chloride being accompanied by rapidly decreasing values of n. The hydrogen chloride added must therefore be involved in some reaction in the gas phase. The only chemical reaction which can be of importance in this connexion, is the removal of chlorine atoms by the triple collision formation of Cl_3 molecules; if this be admitted, then not only chlorine itself but nitrogen and argon also must act in a similar manner, and produce retardations of the main rate of reaction, which will depend on their collision efficiencies as third bodies in the Cl_3 stabilization, and which will increase with the pressure of these added gases.

In support of these conclusions, there is reason to believe, from thermal conductivity measurements incidental to the Budde effect experiments in chlorine (Ritchie & R. L. Smith 1940) that convection effects should be approximately independent of the nature of the added gases, argon, nitrogen and hydrogen chloride. The thermal conductivity of a gas, by kinetic theory, is independent of the pressure of such gas, where pressures are not too small in comparison with the size of the containing vessel, but actual measurement of relative conductivities by means of a heated wire suspended in the reaction vessel showed discrepancies, attributed to convection, which increased with increasing pressure, but were the same for like pressures of all these added gases. Similarly, in the bromine-acetylene reaction referred to, the final rate observed when n approached unity was more or less independent of the nature of the added gas. The only other physical effect which such gases could have here in the hydrogen-chlorine photocombination, is the prevention of diffusion of chain carriers to the walls at high pressures. Calculation shows, however (see later), that for corresponding pressures the rates of diffusion of chlorine atoms, though decreasing at high pressures, are approximately the same for corresponding gas mixtures. This applies also to the diffusion of Cl_3 complexes. If convection were predominant, we should expect the effects of the different gases to be the same; the marked differences in both rates and indices which are observed are therefore taken as indicating that the part played by convection is a minor one. While the possibility of convection effects must be admitted, it would appear that the experimental results above cannot be adequately explained solely on such a basis.

As an alternative, possible effects of oxygen and other inhibitors may be considered. For the experimental conditions of the preceding tables, results were reproducible, Draper effects being sharp and immediate; this, in conjunction with the high quantum efficiencies, suggests that the presence of accidental inhibitors was unlikely. The care taken in the preparation of hydrogen and the added gases, by the use of liquid air and sodium traps, etc., was such as to render the amount of oxygen present extremely small, particularly in the cases of argon, nitrogen and hydrogen. It has already been recognized that the retarding effect of oxygen may be due to the removal of hydrogen and chlorine atoms. The reaction with hydrogen atoms is almost certainly a triple collision process, in which the efficiency of H_2 as a third partner has been shown (Ritchie 1937) to be greater than for A or N_2, although less than for HCl; hence, if the retardation in the presence of nitrogen and argon is due to such a reaction, the oxygen being introduced with the chlorine, a greater retardation for otherwise similar conditions should be apparent on the addition of hydrogen. Experiment shows (table 3), however, that the rate is directly proportional to $[H_2]$ up to pressures of 300 mm., the divergence from linearity at the highest pressure (570 mm.) being only of the order 30 %. In the case of removal of chlorine atoms by oxygen, which may also require a triple collision, it has also been shown (Ritchie 1937) that, for $[H_2]/[Cl_2]$ ratios equal to or greater than unity, retardation of the main action by oxygen takes place predominantly by removal of hydrogen atoms. The effect of Cl atom removal in this way is then negligible, and it is therefore regarded as unlikely that the retardations here in question are due to this cause. It has not been possible to find any expression involving such effects of oxygen which is in agreement with the data, and it would appear that the experimental results again cannot be explained solely on this basis.

It may be here remarked that values of the index n less than 0·5 have been reported for certain conditions by Squire & Allmand (1937), but were attributed by these authors to diffusion of chain carriers from illuminated into dark portions of the reaction mixture. In the present experiments the light beam was so arranged as to be as uniform as possible, and to occupy as much as possible of the reaction vessel volume without actually illuminating the walls parallel to the beam. The low n values obtained are therefore not considered as arising in this way.

To the already established initiation and propagation of reaction chains

$$\begin{array}{lll} Cl_2 \xrightarrow{h\nu} Cl + Cl, & (1), & k_1, \\ Cl + H_2 \rightarrow HCl + H, & (2), & k_2, \\ H + Cl_2 \rightarrow HCl + Cl, & (3), & k_3, \end{array}$$

we therefore add the processes

$$\begin{array}{lll} Cl + Cl_2 + M \rightarrow Cl_3 + M, & (4), & k_4, \\ Cl_3 + M \rightarrow Cl_2 + Cl + M, & (5), & k_5, \\ Cl_3 + Cl_3 \rightarrow 3Cl_2, & (6), & k_6, \\ Cl \rightarrow \text{surface}, & & S_1, \\ Cl_3 \rightarrow \text{surface}. & & S_2. \end{array}$$

Here $S_1[Cl]$ and $S_2[Cl_3]$ represent the number of chlorine atoms and Cl_3 molecules respectively which reach the walls and are removed. In the most general case, S_1 and S_2 will contain factors dependent on diffusion, convection, and the nature of the wall itself. Other chain-ending reactions are possible, but the complexity which they introduce makes it desirable to consider the above scheme first, to find out in what respects and to what extent it fails to account for the results of experiment. The argument is conveniently presented in the following order.

In the stationary state, the following equations hold:

$$\frac{d[Cl]}{dt} = k_1 I_{abs.} - k_2[Cl][H_2] + k_3[H][Cl_2] - k_4[Cl][Cl_2][M] + k_5[M][Cl_3] - S_1[Cl] = 0,$$

$$\frac{d[H]}{dt} = k_2[Cl][H_2] - k_3[H][Cl_2] = 0,$$

$$\frac{d[Cl_3]}{dt} = k_4[Cl][Cl_2][M] - k_5[M][Cl_3] - k_6[Cl_3]^2 - S_2[Cl_3] = 0.$$

Now $$\frac{d[HCl]}{dt} = k_2[Cl][H_2] + k_3[H][Cl_2] = 2k_2[H_2][Cl]$$

and $$\gamma_{HCl} = \frac{1}{I_{abs.}} \frac{d[HCl]}{dt};$$

hence $$\gamma_{HCl} I_{abs.} = \frac{d[HCl]}{dt} = 2k_2[H_2][Cl], \quad \text{or} \quad [Cl] = \frac{\gamma_{HCl} \cdot I_{abs.}}{2k_2[H_2]}.$$

Putting $R = k_4[Cl_2][M]$, $x = I_{abs.}$, $y = \gamma_{HCl}$, we eliminate $[Cl]$ and $[Cl_3]$ and obtain

$$nxy^2 - pxy + qy + rx = K, \qquad (1)$$

where
$$n = \frac{(R+S_1)^2}{(2k_2[\mathrm{H_2}])^2},\quad p = \frac{k_1(R+S_1)}{k_2[\mathrm{H_2}]},$$

$$q = \frac{k_5[M]}{2k_6 k_2[\mathrm{H_2}]}[k_5[M]\,S_1 + (R+S_1)\,S_2],\quad r = k_1^2,\quad K = \frac{k_1}{k_6}k_5[M]\,(k_5[M]+S_2).$$

Neglecting the qy term, we write the remaining terms in the form

$$ny^2 - py - \left(\frac{K}{x} - r\right) = 0;$$

whence
$$y = \frac{p}{2n}\left[1 + \sqrt{\left\{1 + \frac{4nr}{p^2}\left(\frac{K}{rx} - 1\right)\right\}}\right], \tag{2}$$

or
$$\gamma_{\mathrm{HCl}} = \frac{2k_1}{R}\,\frac{k_2[\mathrm{H_2}]}{(1+S_1/R)}\left[1 + \frac{k_5[M]}{\sqrt{(k_1 k_6)}}\,\frac{\sqrt{\{1+S_2/k_5[M]\}}}{\sqrt{I_{\mathrm{abs.}}}}\right]. \tag{3}$$

Since any of the gases of the system may play the part of a stabilizing third body in reaction 4, R ($= k_4[M]\,[\mathrm{Cl_2}]$) must be replaced by

$$[\mathrm{Cl_2}]\,(k_4^{\mathrm{Cl_2}}[\mathrm{Cl_2}] + k_4^M[M]),$$

where M may be $\mathrm{N_2}$, A, HCl and $\mathrm{H_2}$, each with its appropriate value of k_4^M. Similarly $k_5[M]$ is to be replaced by $(k_5^{\mathrm{Cl_2}}[\mathrm{Cl_2}] + k_5^M[M])$.

In considering the experimental trends (*a*)–(*e*) in relation to this equation (3), it is to be remembered that the S_1 and S_2 terms are composite and may contain convection and surface efficiency as well as diffusion factors. If diffusion be the chief factor concerned, S_1 and S_2 will decrease as the pressure increases. Calculation shows (see table 14) that an increase in [HCl], [A], or [$\mathrm{N_2}$] from zero to 400 mm. for [$\mathrm{Cl_2}$] = [$\mathrm{H_2}$] = 43 mm., [HCl] = 15 mm., decreases the rates of diffusion to roughly one-fifth of the former value. On this basis of S_1 and S_2 as determined solely by diffusion, it may be shown that with one exception equation (3) includes the general trends above noted.

As $I_{\mathrm{abs.}}$ increases, S_1 and S_2 will remain constant, and, for constant pressures of reactants, the quantum efficiency will decrease to a constant value as in figure 1. Under such high $I_{\mathrm{abs.}}$ conditions, addition of, say, argon will mean a diminished flow of chlorine atoms to the wall by diffusion; S_1 will decrease and the quantum efficiency will tend to rise. This rise will, however, be succeeded by a decrease by reason of the increase of R, and a maximum may thus appear at an intermediate pressure. This is evident in the argon curves of figure 2. The existence of such a maximum will be dependent on the relative triple collision coefficient in relation to S_1. The rates of diffusion of atoms to the wall are approximately the same for the same pressures of argon, nitrogen and hydrogen chloride; if the surface conditions remain as before, the relative triple collision efficiencies are then, from inspection of figure 2, in the order HCl > $\mathrm{N_2}$ > A. This is in agreement with relative efficiencies quoted for other stabilizing triple collisions; for example, in the HO_2 formation, the efficiencies are respectively 9 : 1·65 : 1 (Ritchie 1937).

A maximum in the quantum efficiency will be observed similarly at the intermediate pressure of chlorine. In the case of hydrogen, a small increase of hydrogen pressure in a mixture which includes 43 mm. of chlorine will not appreciably alter the rate of diffusion of atoms to the wall: the change in S_1 will be relatively small. If hydrogen can act as a stabilizing third molecule, comparison with other triple collisions suggests that such an effect will also be small, on the basis that such a collision efficiency is largely determined by molecular weight. At the lower pressures the quantum efficiency is thus proportional to the hydrogen pressure (table 3), as indicated by the formula. The departure from linearity at the higher hydrogen pressures is in the direction that suggests a small triple collision efficiency, its effect overcoming the expected rise arising from decreased diffusion.

On the addition of further 'inert' gas M, by reason of the now increasing importance of the $I_{\text{abs.}}$ term of the expression, the quantum efficiency will tend to become independent of the pressure and nature of the added gas. This state of affairs will be reached most quickly by the gas of the highest collision efficiencies k_4 and k_5. Such a tendency is evident in the γ-HCl curves of figure 2. At the same time, the quantum efficiencies will become more and more dependent on $I_{\text{abs.}}$; if S_1 and S_2 become vanishingly small, the rate of formation of hydrogen chloride at high values of M will become proportional to the square root of $I_{\text{abs.}}$. The approach to an index n value of 0·5 will thus be most rapid for gases of high k^M efficiency. Since efficiencies for reactions 4 and 5 must be in the same order in respect of the different species of 'inert' molecules, argon will show the least approach to $n = 0{\cdot}5$, which is in agreement with the calculated values of table 11.

Further examination of the formula shows, however, that while the trends (*a*)–(*e*) above may thus be generally accounted for, there is one feature which cannot be so included. The minimum value of n to be obtained by the formula is 0·50, and this is in disagreement with the value 0·39 recorded for 400 mm. hydrogen chloride. The assumption that the nature of the vessel surface is considerably different at 400 mm. from that at 15 mm. by reason of hydrogen chloride adsorption, cannot yield a value less than 0·50, since comparison is made of results at 400 mm. at the two intensities. To account for the figure 0·39 by formula (3), S_1 would need to be much greater for the high $I_{\text{abs.}}$ conditions; in other words, we would require here the effect of convection to explain it.

Neglecting this divergence for the present, we may inquire if the other results can be numerically accounted for by the formula. It has already been shown that the quantum efficiency is directly proportional to the hydrogen pressure at the lower pressures. Higher pressures are referred to later. Consideration of the effect of altered chlorine concentration is more complex, because increase of $[Cl_2]$ means in general an altered $I_{abs.}$, which in the experiments of tables 6 and 7 has not been exactly compensated by reducing the incident light intensity. By reason also of the photometric method of measuring the rate of reaction, accurate results could not be obtained for a change of 15 mm. hydrogen chloride as before. More important is the fact that it is difficult to calculate exact values of S_1 and S_2, even relatively, on a purely diffusion basis, since alteration in $[Cl_2]$ alters the distribution of chlorine atoms initially produced by the light beam. At the higher chlorine pressures however, diffusion will be small and the effect of all these factors minimized. Since chlorine must be taken as itself playing the part of an M molecule, formula (3) indicates that at high $[Cl_2]$, the quantum efficiency should be inversely proportional to the chlorine pressure. The experimental results of table 7 support this conclusion: in the final column of table 12 are given values of $\gamma[Cl_2]$, which will be observed to be approximately constant at higher pressures. It is of further interest to note that if the $I_{abs.}$ term of formula (3) were to become small (as at high values of $I_{abs.}$), the quantum efficiency would tend to become inversely proportional to the square of the chlorine pressure. Although no experiments were carried out here from this point of view, Squire & Allmand (1937) did observe such a tendency at certain high chlorine pressures.

TABLE 12. LARGE REACTION VESSEL

Volume 178·0 c.c.; $[H_2] = 43{\cdot}0$ mm.

$[Cl_2]$ mm.	γ_{HCl}	$\gamma_{HCl}[Cl_2] \times 10^{-4}$
67·2	39,500	266
103·7	33,900	352
204·6	18,000	366
308·9	12,100	372

In considering the experimental series in which the pressures of added gases are the only variables, it is first to be noted that the low index n values at high hydrogen chloride pressures have already been shown to be in disagreement with formula (3), if convection effects are absent. In spite of this it has been found convenient to consider all such series from the viewpoint that S_1 and S_2 are determined solely by the rates of diffusion of the chlorine atoms and Cl_2 molecules respectively and by the ability of the vessel walls to remove by absorption these intermediates on collision.

The relative rate of diffusion of a particle A through a gas X is given by the expression

$$D_X = \sqrt{\left(\frac{1}{M_A}+\frac{1}{M_X}\right)}\,\sigma_{AX}^{-2}[X]^{-1},$$

where M_A = molecular weight of diffusing particle,

M_X = molecular weight of gas X,

σ_{AX} = sum of radii of diffusing particle and gas molecule X,

and $[X]$ = pressure of gas X.

For a mixture of chlorine with other gases (M_1, M_2, etc.) the time of diffusion through a given distance is taken as

$$t = \frac{1}{D_{\mathrm{Cl}_2}} + \frac{1}{D_{M_1}} + \frac{1}{D_{M_2}} + \dots,$$

and the relative rate of diffusion is then

$$D = \left(\frac{1}{D_{\mathrm{Cl}_2}} + \frac{1}{D_{M_1}} + \frac{1}{D_{M_2}} + \dots\right)^{-1}.$$

This formula may be applied when the relative distribution of diffusing particles remains unaltered throughout the reaction vessel.

Values of D_1 and D_2 for chlorine atoms and Cl_3 molecules respectively were calculated from the following data:

$$\sigma_{\mathrm{Cl}_2} = 3{\cdot}6 \times 10^{-8}\,\mathrm{cm.},\quad \sigma_{\mathrm{H}_2} = 2{\cdot}3 \times 10^{-8}\,\mathrm{cm.},$$

$$\sigma_{\mathrm{HCl}} = 3{\cdot}0 \times 10^{-8}\,\mathrm{cm.},\quad \sigma_{\mathrm{N}_2} = 3{\cdot}1 \times 10^{-8}\,\mathrm{cm.},\quad \sigma_{\mathrm{A}} = 2{\cdot}86 \times 10^{-8}\,\mathrm{cm.}$$

The diameter of the chlorine atom was taken as $3{\cdot}0 \times 10^{-8}$ cm., and that of the Cl_3 molecule as $6{\cdot}0 \times 10^{-8}$ cm. Concentrations $[X]$ were expressed in millimetres.

In the absence of convection, diffusion and surface removal of chain carriers may then be combined in the form

$$S = mD,$$

where m, in addition to accounting for wall efficiency, includes factors depending on the temperature, the size and shape of the reaction vessel and the position of the light beam; m will be constant if all these factors remain the same. The constancy of wall efficiency is the only one which cannot be directly controlled; if this efficiency is not unity, then in the general case m will be different for chlorine atoms and for Cl_3 molecules and will vary with the nature of the surface and the amount of each different gas adsorbed on it. Hartley, Henry & Whytlaw-Gray (1939) have shown that the adsorption of gases on silica at ordinary temperatures and pressures, though much less than that on soft glass, is by no means negligible; it increases with the pressure and for equal pressures of argon, nitrogen and hydrogen chloride rises in that order. The adsorption of the more easily condensable gases might then be expected to produce the maximum change. Craggs & Allmand (1937) have shown, however, that the addition of water vapour only slightly reduces the rate of hydrogen chloride photosynthesis and similarly Budde effects in chlorine (Ritchie & R. L. Smith 1940), and in bromine (W. Smith, Ritchie & Ludlam 1937), were only slightly reduced in vessels of soft glass saturated with the water vapour under conditions where a considerable portion of the atoms were removed by surface action. The changes in the surface efficiencies were therefore small, and it might thus be argued that in the present experiments with the less easily condensable gases argon and nitrogen such effects would be similarly unimportant, especially as an appreciable pressure of chlorine was already present. The previous examinations of the chlorine and bromine Budde effects do not contradict such assumptions.

Formula (3) is then applied in the form

$$\gamma_{\mathrm{HCl}} = \frac{a'[\mathrm{H_2}]}{R'\left(1+\frac{m_1 D_1}{R'}\right)}\left[1+\frac{b'M'\sqrt{\left(1+\frac{m_2 D_2}{M'}\right)}}{\sqrt{I_{\mathrm{abs.}}}}\right], \tag{4}$$

where
$$a' = \frac{2k_1 k_2}{k_4^{\mathrm{Cl_2}}}, \qquad R' = [\mathrm{Cl_2}]([\mathrm{Cl_2}]+k_M[M]).$$

$$b' = \frac{k_5^{\mathrm{Cl_2}}}{\sqrt{(k_1 k_6)}}, \quad M' = [\mathrm{Cl_2}]+k_M[M].$$

Since relative M efficiencies must be the same for reactions 4 and 5,

$$\frac{k_4^{\mathrm{Cl_2}}}{k_5^{\mathrm{Cl_2}}} = \frac{k_4^{\mathrm{A}}}{k_5^{\mathrm{A}}} = \frac{k_4^{\mathrm{N_2}}}{k_5^{\mathrm{N_2}}} = \frac{k_4^{\mathrm{HCl}}}{k_5^{\mathrm{HCl}}} = \frac{k_4^{\mathrm{H_2}}}{k_5^{\mathrm{H_2}}};$$

the coefficient k_M is thus
$$\frac{k_4^{\mathrm{M}}}{k_4^{\mathrm{Cl_2}}} = \frac{k_5^{\mathrm{M}}}{k_5^{\mathrm{Cl_2}}}.$$

The constants m_1 and m_2 contain $1/k_4^{\mathrm{Cl_2}}$ and $1/k_5^{\mathrm{Cl_2}}$ respectively.

With the exception of three series, viz. the argon, nitrogen and hydrogen chloride curves at high intensities, the equation (4) with suitable constants adequately accounts for the experimental data. The calculations are outlined in tables 13–14: the factor

$$\left(1+\frac{b'M'\sqrt{\left(1+\frac{m_2 D_2}{M'}\right)}}{\sqrt{I_{\mathrm{abs.}}}}\right)$$

is denoted by W. It will be observed that the only essential variable is k_M. The values of the constants used were:

$$a' = 6{\cdot}36\times10^{6}, \quad b' = 1{\cdot}80\times10^{3}, \quad m_1 = 5400\times10^{-12},$$

$$m_2 = 1100\times10^{-12}, \quad k_{\mathrm{H_2}} = 1{\cdot}15, \quad k_{\mathrm{HCl}} = 6{\cdot}05, \quad k_{\mathrm{A}} = 1{\cdot}00, \quad k_{\mathrm{N_2}} = 1{\cdot}60.$$

For convenience in calculation, the experimental values in the added gas series have been taken from the smoothed curves. $I_{\text{abs.}}$ represents the average total quanta per second absorbed throughout the reaction vessel.

In view of the slightly lower experimental results in the nitrogen series for $[N_2] = 0$ (see table 10*b*), it is probable that the S_1 term here is slightly greater than that for the other series. The ratios of the final columns of the above tables show a satisfactory agreement between theory and experiment.

Where $I_{\text{abs.}}$ is the only variable, the concentrations of reactants being steady, S_1 and S_2 will be constant if convection is absent and formula (4) reduces to

$$\gamma_{\text{HCl}} = a\left[1 + \frac{b}{\sqrt{I_{\text{abs.}}}}\right]. \tag{5}$$

The results of table 13 are in effect given by taking $a = 11{,}200$, $b = 0{\cdot}91 \times 10^6$; over the range of the 500 : 1 variation of $I_{\text{abs.}}$, the maximum divergence is 8 % and the mean deviation 1 %, which is well within the experimental error. No other expression has been found to give anything like such agreement, and it is therefore concluded that the neglect of the qy term of the general equation and the treatment of S_1 and S_2 as constant are justified for these conditions. It is then to be expected that a formula of similar type should apply to the parallel series of results for the large reaction vessel (table 5). The vessels were of similar shape and material; if the walls have similar efficiencies of removal of chain carriers as before, both S_1 and S_2 should be smaller in the large vessel for similar concentrations of reactants; a should be larger and b smaller. Experimental results for the large vessel are in fact satisfactorily represented by formula 5, with $a = 31{,}600$ and

TABLE 13. VARIATION OF $I_{\text{abs.}}$

Vessel volume = 31·3 c.c.; $[\text{Cl}_2] = [\text{H}_2] = 43{\cdot}0$ mm.; $[\text{HCl}] = 15{\cdot}0$ mm.; $m_1 D_1 = 16{,}500$; $m_2 D_2 = 1210$.

λ A	$I_{\text{abs.}} \times 10^{-12}$ quanta/sec.	$\gamma_{\text{HCl}} \times 10^{-4}$ calc.	$\gamma_{\text{HCl}} \times 10^{-4}$ exp.	$\gamma_{\text{calc.}}/\gamma_{\text{exp}}$
3650	153·4	1·19	1·18	1·01
3650	149·0	1·195	1·205	0·99
3650	78·5	1·23	1·16	1·05
3650	25·4	1·315	1·21	1·08
3650	21·4	1·33	1·28	1·04
3650	17·85	1·35	1·305	1·04
4060	15·8	1·365	1·28	1·07
4060	10·85	1·42	1·39	1·02
3650	8·10	1·47	1·42	1·03
4060	5·29	1·55	1·60	0·97
3650	4·36	1·60	1·60	1·00
4060	3·25	1·67	1·73	0·97
3650	2·2	1·795	1·78	1·01
4060	1·47	1·95	2·06	0·95
4060	0·855	2·21	2·205	1·00
4060	0·306	2·945	2·73	1·08

TABLE 14

Vessel volume = 31·3 c.c.; $[Cl_2]$ = 43·0 mm.

	$[M']$	m_1D_1	m_2D_2	W	$\gamma_{calc.} \times 10^{-4}$	$\gamma_{exp.} \times 10^{-4}$	$\gamma_{calc.}/\gamma_{exp.}$
Variation of hydrogen. [HCl] = 15·0 mm.; av. $I_{abs.} = 11·5 \times 10^{13}$							
$[H_2]$							
43	183·2	16,500	1210	1·27	1·42	1·43	0·99
100	248·7	14,050	1056	1·30	3·36	3·40	0·99
200	353·7	11,200	855	1·35	6·50	6·50	1·00
300	478·7	9,300	720	1·40	8·85	9·0	0·98
400	593·7	7,900	620	1·45	11·1	11·0	1·01
500	708·7	6,870	547	1·50	12·8	12·6	1·02
600	823·7	6,100	488	1·55	14·4	14·2	1·01
Variation of hydrogen chloride. $[H_2]$ = 43·0 mm.; av. $I_{abs.} = 16·8 \times 10^{13}$							
[HCl]							
15	183·2	16,500	1210	1·22	1·37	1·38	1·00
100	697·5	8,300	610	1·42	1·01	0·97	1·04
200	1302	5,230	384	1·65	0·737	0·73	1·01
300	1907	3,820	280	1·89	0·603	0·60	1·00
400	2512	2,980	220	2·15	0·53	0·54	0·98
Variation of argon. $[H_2]$ = 43·0 mm.; [HCl] = 15·0 mm.; av. $I_{abs.} = 16·7 \times 10^{13}$							
[A]							
0	183·2	16,500	1210	1·22	1·37	1·33	1·03
100	283·2	7,750	561	1·22	1·65	1·60	1·03
200	383·2	5,020	363	1·24	1·59	1·60	1·00
300	483·2	3,760	270	1·27	1·42	1·44	0·99
400	583·2	2,960	214	1·30	1·27	1·26	1·01
Variation of nitrogen. $[H_2]$ = 43·0 mm.; [HCl] = 15·0 mm.; av. $I_{abs.} = 15·9 \times 10^{13}$							
$[N_2]$							
0	183·2	16,500	1210	1·23	1·38	1·29	1·07
100	343·2	7,800	580	1·25	1·52	1·47	1·03
200	503·2	5,070	385	1·30	1·34	1·42	0·94
300	663·2	3,780	276	1·36	1·16	1·16	1·00
400	823·2	3,000	230	1·42	1·01	0·96	1·05

$b = 0·85 \times 10^6$, as shown in table 15, where again $I_{abs.}$ is expressed as total quanta absorbed per second throughout the vessel.

For correct comparison of the b values for the two vessels, $I_{abs.}$ must be expressed in some common unit. On converting to quanta/sec./c.c., b for the small vessel is $0·163 \times 10^6$ and for the large vessel is $0·065 \times 10^6$. The values of a and b are thus in accordance with expectation. Further, the limiting value of γ_{HCl} in the small vessel ($I_{abs.} = \infty$) is

$$a'[H_2]/R'\left(1+\frac{m_1D_1}{R'}\right) = 11{,}200$$

for $[Cl_2] = [H_2] = 43$, [HCl] = 15 mm.: if m_1D_1 is negligible, the resultant value of 35,000 (since $1+\frac{m_1D_1}{R} = 3·1$) must be greater than the corresponding limiting value in the large vessel, for which diffusion must still

TABLE 15. VARIATION OF $I_{\text{abs.}}$

Vessel volume = 178·0 c.c.; $a = 31{,}600$; $b = 0{\cdot}85 \times 10^6$;
$[H_2] = [Cl_2] = 43{\cdot}0$ mm.; $[HCl] = 15$ mm.; $\lambda = 4060$ A

$I_{\text{abs.}} \times 10^{-13}$ quanta/sec.	$\gamma_{HCl} \times 10^{-4}$ calc.	$\gamma_{HCl} \times 10^{-4}$ exp.	$\gamma_{\text{calc.}}/\gamma_{\text{exp.}}$
118·0	3·40	3·34	1·02
117·0	3·41	3·42	1·00
112·8	3·41	3·55	0·96
81·7	3·46	3·40	1·02
79·0	3·46	3·54	0·98
34·1	3·62	3·44	1·05
16·23	3·81	3·54	1·08
8·99	4·06	3·99	1·02
4·98	4·37	4·65	0·94

exist although to a less degree. The limiting experimental efficiency in the large vessel is 31,600, some residual surface action being thus indicated. If the relative efficiency of removal of Cl and Cl_3 carriers is the same in both vessels, we should expect the $1/\sqrt{I_{\text{abs.}}}$ coefficient to decrease to 0·415 times its former value; the actual ratio, from the figures already given, is

$$0{\cdot}065/0{\cdot}163 = 0{\cdot}40.$$

The relative triple collision coefficients, k_{N_2} being taken as unity, are

$$k_{HCl} : k_{N_2} : k_{H_2} : k_A : k_{Cl_2} = 3{\cdot}8 : 1{\cdot}0 : 0{\cdot}72 : 0{\cdot}63 : 0{\cdot}63.$$

In the HO_2 stabilization corresponding ratios are

$$k_{HCl} : k_{N_2} : k_A = 5{\cdot}4 : 1{\cdot}0 : 0{\cdot}61,$$

while the $k_{N_2} : k_A$ ratios for the bromine, iodine and ozone stabilizations are again all of the order 1·0 : 0·60 (see Ritchie 1937). In the case of hydrogen the high ratio $k_{H_2} : k_{N_2}$ of 1·35 for the $H + O_2 + H_2$ reaction is probably due to the fact that this is not a stabilization reaction but one which yields products other than HO_2; where the hydrogen molecule acts solely as a stabilizer as in iodine and bromine atom recombinations the efficiency is of the order of that of argon, as above. Similarly, in the case of chlorine, the high value of $k_{Cl_2} : k_{N_2} = 3{\cdot}57$ for the $H + O_2 + Cl_2$ reaction is probably due to the production of HCl and ClO_2 rather than HO_2. On the other hand, Griffiths & Norrish (1931) reported a ratio of 2·24 : 1 for the reaction $Cl + NCl_3 + M \rightarrow NCl_4 + M$, and this is apparent disagreement with the present figure. It is, however, to be noted that when a molecule may play the part of reactant as well as stabilizer, as in the reactions

$$H + O_2 + O_2 \rightarrow HO_2 + O_2 \quad \text{and} \quad O + O_2 + O_2 \rightarrow O_3 + O_2,$$

the derived value for $k_{O_2} : k_{N_2}$ is then approximately one-third of its value for a reaction such as $Br + Br + M \rightarrow Br_2 + M$. A ratio one-third of 2·24, viz. 0·75, might be expected for the reaction $Cl + Cl_2 + Cl_2 \rightarrow Cl_3 + Cl_2$ and this is more in accordance with the calculations of the tables above. The relative triple collision efficiencies so found may thus be said to be in fair agreement with those observed in other stabilizations.

An estimate may be made of the ratio k_4/k_5 from the constants

$$a' = 2k_1 k_2/k_4 \quad \text{and} \quad b' = \frac{k_5}{\sqrt{(k_1 k_6)}}$$

as used above. The velocity coefficient k_2 has been calculated on the basis of a collision efficiency $3{\cdot}5 \times 10^{-5}$ at 25° (Rodebush & Klingelhoefer 1933), that of reaction 6 being taken as unity. The molecular diameters of the diffusion calculations were employed. Taking $k_1 = 2$, expressing all concentrations in mm., and correcting for the fact that the b' value is related above to $I_{\text{abs.}}$ expressed as total quanta absorbed per second throughout the reaction vessel of 31·3 c.c. volume, we find $k_4/k_5 = 0{\cdot}059$ mm.$^{-1}$. This is to be compared with 0·0049 obtained by Craggs *et al.* (1937) and 0·78 calculated by Rollefson & Eyring (1932). To bring our figure to agreement with that of Rollefson & Eyring a collision efficiency of approximately 0·057, i.e. 1 in 175, would be required for reaction 6.

It has already been pointed out that the factors m_1 and m_2 contain $1/k_4^{Cl_3}$ and $1/k_5^{Cl_3}$ respectively. From the ratio k_4/k_5 above, it then appears that the vessel surface had an efficiency of removal of Cl_3 complexes by adsorption roughly three times that for chlorine atoms.

According to the general Cl_3 mechanism above, the limiting value of the quantum efficiency at high $I_{\text{abs.}}$ in the small vessel is $\dfrac{2k_1 k_2[H_2]}{R(1+S_1/R)} = 11{,}200$, for $[Cl_2] = [H_2] = 43$ mm., $[HCl] = 15$ mm. With $k_1 = 2$, representing two chlorine atoms produced per quantum absorbed, this is approximately equal to $\dfrac{k_2[H_2]}{3{\cdot}3k_4^{Cl_3}[Cl_2]\,[Cl_2]}$ since $R = k_4^{Cl_3}[Cl_2]\,[M]$ and $M' = 183$, $1 + m_1 D_1/R' = 3{\cdot}1$. Thus $\dfrac{k_2[H_2]}{k_4^{Cl_3}[Cl_2]\,.\,[Cl_2]} = 3{\cdot}7 \times 10^4$ for these conditions. From consideration of the reactions $Cl + H_2 \rightarrow HCl + H$ and $Cl + Cl_2 + Cl_2 \rightarrow Cl_3 + Cl_2$, it is obvious that $k_2[H_2]/k_4^{Cl_3}[Cl_2]\,[Cl_2]$ is the ratio of effective binary collisions of Cl atoms to effective triple collisions. The ratio of total binary to total ternary collisions at such pressures is of the order 10^4, and if e_2 and e_4 are the respective collision efficiencies, the ratio of effective binary to effective ternary collisions will be $(e_2/e_4) \times 10^4$, which is then equal to $3{\cdot}7 \times 10^4$. The collision efficiency e_2 is of the order $3{\cdot}5 \times 10^{-5}$ at 25°; hence $0{\cdot}35/e_4 = 3{\cdot}7 \times 10^4$ or the efficiency of the ternary collision $Cl + Cl_2 + Cl_2 \rightarrow Cl_3 + Cl_2$ is of the order 10^{-5}, an unexpectedly small value as compared with other triple collisions. In explanation, it might be argued that we deal here with the formation of a complex which is itself not a normal stable molecule, and that normal values would not be expected. The low efficiency is then to be allied to the very high quantum yields which have been recorded for the hydrogen chlorine reaction. If there is no evidence for the existence of Cl_3 molecules from the study of other photochemical reactions involving chlorine, it may be that it is only in reactions with very high quantum yields, where other chain-ending processes operate to a small extent only, that the Cl_3 mechanism will become appreciable.

TABLE 16. SMALL REACTION VESSEL

Volume = 31·3 c.c.; $[Cl_2] = [H_2] = 43·0$ mm.

[A] mm.	$\gamma_{calc.} \times 10^{-4}$	$\gamma_{exp.} \times 10^{-4}$	$\gamma_{calc.}/\gamma_{exp.}$
	Argon, $I_{abs.} = 149 \times 10^{12}$		
0	1·21	1·21	1·00
100	1·45	1·29	1·12
200	1·38	1·22	1·13
300	1·22	1·03	1·18
400	1·07	0·82	1·30
	Nitrogen, $I_{abs.} = 139 \times 10^{12}$		
$[N_2]$ mm.			
0	1·05	1·05	1·00
100	1·23	1·01	1·22
200	1·09	0·84	1·30
300	0·93	0·66	1·41
400	0·80	0·48	1·67
	Hydrogen chloride, $I_{abs.} = 153 \times 10^{12}$		
[HCl] mm.			
15	1·20	1·21	0·99
100	0·81	0·63	1·29
200	0·54	0·30	1·81
300	0·41	0·16	2·60
400	0·34	0·14	2·45

There remain for consideration three series, the nitrogen, argon and hydrogen chloride curves at high intensities. The general mechanism above has accounted satisfactorily for a 500-fold variation in $I_{abs.}$ at low pressures and for the inert gas curves at low intensities and high pressures, on the basis that diffusion alone accounts for the variation in S. It does not follow that a similar state of affairs exists at high intensities in conjunction with high pressures; convection, other chain endings, e.g. $Cl + Cl_3 + 2Cl_2$, and the effect of 'impurities' may all be expected to have a greater effect at the higher intensities. Application of the formula 4 to the three remaining series is given in table 16. The necessary constants have been taken as before, except in the case of the nitrogen series where m_1 and m_2 have been increased to 6700×10^{-12} and 1350×10^{-12} respectively in view of the fact that the quantum efficiency at $N_2 = 0$ is some 12 % less than the usual value at the high intensity.

It has already been pointed out that index values n of less than 0·5, as calculated from the hydrogen chloride curves at high pressures, cannot be accounted for by formula (3). The discrepancies of table 16 indicate that some other factor or factors become of increasing importance at high intensities and high pressures not only in the hydrogen chloride series but in the nitrogen and argon series as well, though to a less degree. Detailed consideration of other possible values of m, etc., and of the previously neglected y term of the general formula has not been of assistance in removing these discrepancies. While these departures from the general scheme above have not been consistently accounted for, the following considerations may be briefly presented.

On the basis of the thermal conductivity measurements previously referred to, convection effects for a given heat production will be the same for corresponding pressures of all the added gases argon, nitrogen and hydrogen chloride, and therefore approximately the same departures from the general reaction mechanism will occur if convection is the factor concerned. The discrepancy is much greater for hydrogen chloride, where such discrepancy cannot be explained by altered surface efficiency of removal of chain carriers on the basis of the formula used. This is the more remarkable in that the rate of reaction at high [HCl] was some five times less than the corresponding rate with argon and the possible convection will be correspondingly less.

The efficiency of the removal of hydrogen atoms by the triple collision process $H + O_2 + M$ is nine times as efficient for hydrogen chloride as for argon. Of the gases concerned, it is believed that chlorine is the most difficult to render free from oxygen. Since the hydrogen chloride used was prepared from this chlorine by illumination, the conditions of high [HCl] are those in which the highest oxygen concentrations may be expected. There will then be competition between two processes, one involving $I^{\frac{1}{2}}_{\text{abs.}}$ and the other involving $I^{1}_{\text{abs.}}$; at high intensities the latter, removing hydrogen atoms, will be more appreciable, and might cause the abnormally low indices observed. That the rate of photosynthesis may be very sensitive to traces of oxygen has already been indicated by the low efficiency of the $Cl + Cl_2 + Cl_2$ reaction. One feature against such an explanation is that the hydrogen chloride curves at high [HCl] tend to approach limiting constant values, and this would not be the case if oxygen is being progressively introduced as [HCl] rises.

The chain ending $Cl + Cl_3 \rightarrow 2Cl_2, \ldots, k_7$ introduces a factor which does act in the required direction. By the method previously adopted, with

$$R = k_4[Cl_2][M], \quad x = I_{\text{abs.}}, \quad y = \gamma_{HCl},$$

the following equation may be derived:

$$mx^2y^2 - lxy^3 + nxy^2 - pxy + qy + rx = K. \tag{6}$$

By neglecting the m, l and q terms, the remaining terms may be written as before in the form

$$ny^2 - py - \left(\frac{K}{x} - r\right) = 0;$$

whence

$$y = \frac{p}{2n}\left[1 + \sqrt{\left\{1 + \frac{4nr}{p^2}\left(\frac{K}{rx} - 1\right)\right\}}\right], \tag{7}$$

where, now,

$$\frac{p}{2n} = \frac{k_1 k_2 \mathrm{H}_2}{(R+S_1)} \frac{\left[2 - \dfrac{k_7}{k_6}\dfrac{S_2}{(R+S_1)}\right]}{\left[1 + \dfrac{2k_5 M k_7 R}{k_6(R+S_1)^2} - \dfrac{k_7}{k_6}\dfrac{S_2}{(R+S_1)}\right]},$$

$$\frac{4nr}{p^2} = 4\frac{\left[1 + \dfrac{2k_5 M k_7 R}{k_6(R+S_1)^2} - \dfrac{k_7 S_2}{k_6(R+S_1)}\right]}{\left[2 - \dfrac{k_7}{k_6}\dfrac{S_2}{(R+S_1)}\right]^2},$$

and

$$\frac{K}{r} = \frac{k_5 M(k_5 M + S_2)}{k_1 k_6}.$$

Equation (7) reduces to the previous equation (2) if $4nr/p^2 = 1$, a value indicated by the expression for $4nr/p^2$ above if the terms involving k_7 are negligible. It would appear therefore that for the low-pressure conditions such terms are inappreciable. At high M values $\frac{k_7 S_2}{k_6(R+S_1)}$ will tend to become smaller but $\frac{2k_5 M k_7 R}{k_6(R+S_1)^2}$ and $\frac{K}{r}$ will increase; calculation then applied to the ratio of quantum efficiencies at 400 mm. hydrogen chloride shows that the 'abnormal' index $n = 0{\cdot}39$ may be accounted for in this way. If, for example, $4nr/p^2 = 2{\cdot}25$ and $K/r = 85 \times 10^{12}$, to be compared with low-pressure values ([HCl] = 15 mm.) of 1·00 and $0{\cdot}846 \times 10^{12}$, the required ratio of quantum efficiencies at the two $I_{\mathrm{abs.}}$ values may be obtained. Further calculation on these lines, however, introduces difficulties. Detailed consideration of other possible values of m, etc., as well as of the terms previously regarded as negligible, has not produced calculated values in reasonable agreement with experiment throughout. It has to be noted, however, that if the correcting term $\frac{2k_5 M k_7 R}{k_6(R+S_1)^2}$ is of the required order above, i.e. 1·25 for the high-pressure conditions [HCl] = 400 mm., then

$$\frac{2k_5 M k_7 R}{k_6(R+S_1)^2} = \frac{2k_5 k_7}{43k_6 k_4} \bigg/ (1 + m_1 D_1/R') = \frac{k_7}{k_6}\frac{2}{43 \times 0{\cdot}059 \times 1{\cdot}06} = 1{\cdot}25,$$

whence $k_7 = 1{\cdot}7k_6$. The coefficient k_7 is thus of the same order of magnitude as k_6, as would be expected on general grounds. In other words, it is thus probable that the correcting term is of real significance.

The reaction $$\mathrm{Cl} + \mathrm{Cl} + M \rightarrow \mathrm{Cl}_2 + M \tag{8}$$

again introduces a similarly acting factor, the term

$$\frac{k_8 M k_5 M (k_5 M + S_2)}{k_6 (R + S_1)^2} \quad \text{replacing} \quad \frac{2k_5 M k_7 R}{k_6 (R + S_1)^2}.$$

Calculation again indicates that if reaction (8) has unit efficiency such a chain ending will tend to become appreciable at the higher pressures and intensities, but again calculated values in satisfactory agreement with experimental data have not been obtained throughout the three series.

In conclusion, it may be said that in spite of the numerical discrepancies discussed above, the Cl_3 hypothesis does account generally for the observed features of this complicated reaction. The quantum efficiency is normally proportional to the hydrogen pressure. When surface action is present, we may expect an increase in quantum efficiency as the chlorine pressure rises, or a region where efficiency is independent of chlorine pressure or a region where the rate is strongly retarded by chlorine. Similar remarks apply to nitrogen, argon and hydrogen chloride. With hydrogen chloride an initial rise has here not been observed with increasing pressure, but is to be expected with greater surface action. The index values of n, expressing the effects of altered intensity of absorbed light, may range from unity to values less than 0·5. Finally, the relative triple collision efficiencies of hydrogen, argon, nitrogen and hydrogen chloride in the stabilizing collision $\mathrm{Cl} + \mathrm{Cl}_2 + M$ are here found to be of the same order as in other stabilizing triple collisions. Experiments on the addition of otherwise inert gases do not appear to have been previously reported; the present observations on the tendency of the index n to approach unity at high light intensities are probably due to the fact that the light absorbed was here of the high order $0{\cdot}5$ to $7{\cdot}5 \times 10^{12}$ quanta/sec./c.c. at the relatively low chlorine pressure of 43 mm. Bodenstein & Winter (1936) used intensities of the order 10^9 quanta/sec./c.c. at 375 mm. chlorine, while those of Craggs, Squire & Allmand (1937) were of the order 10^{10}–10^{13} at 166 and 450 mm. The general tendencies noted by these latter authors have been confirmed in the present experiments.

The authors wish to express their gratitude to the Carnegie Trustees for a Teaching Fellowship (M. R.) and Scholarship (D. T.) and to thank the Trustees of the Moray Research Fund for a grant for apparatus.

References

Bodenstein, M., Brenschede, W. & Schumacher, H. J. 1938 *Z. phys. Chem.* **B, 40, 125.**
Bodenstein, M. & Unger, W. 1931 *Z. phys. Chem.* B, **11**, 268.
Bodenstein, M. & Winter, R. 1936 *S.B. preuss. Akad. Wiss.* p. 1.
Craggs, H. C. & Allmand, A. J. 1937 *J. Chem. Soc.* p. 1889.
Craggs, H. C., Squire, G. V. V. & Allmand, A. J. 1937 *J. Chem. Soc.* p. 1878.

Cremer, E. 1927 *Z. phys. Chem.* **128**, 285.
Göhring, R. 1921 *Z. Elektrochem.* **27**, 511.
Griffiths, J. G. A. & Norrish, R. G. W. 1931 *Trans. Faraday Soc.* **27**, 451.
Hartley, G. A. R., Henry, T. H. H. & Whytlaw-Gray, R. 1939 *Trans. Faraday Soc.* **35**, 1452.
Müller, K. L. & Schumacher, H. J. 1938 *Z. phys. Chem.* B, **39**, 352.
Norrish, R. G. W. & Ritchie, M. 1933 *Proc. Roy. Soc.* A, **140**, 713.
Potts, J. C. & Rollefson, G. K. 1935 *J. Amer. Chem. Soc.* **57**, 1027.
Ritchie, M. 1937 *J. Chem. Soc.* p. 857.
Ritchie, M. & Norrish, R. G. W. 1932 *Nature, Lond.*, **129**, 243.
Ritchie, M. & Norrish, R. G. W. 1933 *Proc. Roy. Soc.* A, **140**, 99, 112.
Ritchie, M. & Smith, R. L. 1940 *J. Chem. Soc.* p. 394.
Rodebush, W. H. & Klingelhoefer, W. C. 1933 *J. Amer. Chem. Soc.* **55**, 130.
Rollefson, G. K. & Eyring, H. 1932 *J. Amer. Chem. Soc.* **54**, 170.
Smith, W., Ritchie, M. & Ludlam, E. B. 1937 *J. Chem. Soc.* p. 1680.
Squire, G. V. V. & Allmand, A. J. 1937 *J. Chem. Soc.* p. 1869.

Crystal theory of metals: calculation of the elastic constants

By K. Fuchs† and H. W. Peng
University of Edinburgh

(*Communicated by M. Born, F.R.S.—Received* 25 *September* 1941)

The approximate equations of motion for the electrons in a cyclic lattice of a metal are set up with the help of the self-consistent field. The displacements of the ions are then considered as perturbations of the motion of the electrons. The change of the boundary is compensated by a co-ordinate transformation. The change of the potential energy of the lattice due to a homogeneous deformation is calculated by the perturbation method. The calculated values of the elastic constants are found to be in satisfactory agreement with the observed values.

Introduction

The character of metallic binding has been revealed by the calculations of Wigner & Seitz (1933, 1934). But little progress has since been made in developing a crystal theory of metals comparable with the classical crystal theory of ionic lattices of Born (1923), apart from the calculation of the elastic properties (Fröhlich 1937; Fuchs 1936*a*; Bardeen 1938).

Recently one of us (Fuchs 1940) suggested that the method of the operator calculus is suitable for further progress in this direction. In this paper we wish to attempt such a development, and the general method of treatment is outlined. The results are applied to a calculation of the elastic constants. The calculation of the thermal frequencies will be given in a later paper by one of us.

---

† Now at the University of Birmingham.

The abstract notation of the operator calculus of metals, though well suited for explaining the ideas involved, is rather unwieldy, if applied to more complicated problems (e.g. the thermal frequencies). For this reason we shall use in this paper the equivalent methods of the wave theory.

To avoid getting involved in a great number of numerical calculations, it is eventually necessary to introduce the simple Wigner-Seitz approximation for the calculation of the wave functions. Yet it is worth while to start with a more general case. Not only does it give a correct picture of the approximations involved; it also allows us to apply those corrections to the simple Wigner-Seitz approximation, which have been studied by Wigner & Seitz in their succeeding papers, e.g. the lattice energy and the correlation energy between the electrons.

We shall therefore start from the Fock-Dirac system of equations for the wave functions of the metallic electrons, which has recently been studied by one of us (Peng 1941, in the following quoted as I). The perturbation theory developed in that paper can be immediately employed for our present purposes.

In § 1, the Fock-Dirac equations for the electrons in a cyclic lattice of metals are set up, and the general expression for the lattice potential energy is given. In § 2, the ionic displacements are treated as perturbations and the perturbation method is described. § 3 is devoted to the explicit calculation of certain perturbation operators and to the derivation of certain useful relations expressing them in terms of a substitution operator. We then confine ourselves to the case of homogeneous deformation and enter into the detailed calculation of the elastic constants.

The results for the elastic constants show that, as far as the volume independent combinations of the elastic constants are concerned, the approximations introduced in previous calculations are justified. But for the compressibility an improved equation is derived, which compares favourably with experimental results.

1. FOCK-DIRAC EQUATIONS FOR THE ELECTRONS IN METALS

We consider an infinite crystal with lattice vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$. The lattice points are then given by $\mathbf{a}^l = l_1\mathbf{a}_1 + l_2\mathbf{a}_2 + l_3\mathbf{a}_3$. In the equilibrium configuration each lattice point is occupied by an ion. But if we consider homogeneous deformations or thermal oscillations, the ions will be displaced. We may then express the position of the lth ion as a function of $\mathbf{a}^l$

$$\mathbf{r}^l = \mathbf{r}(\mathbf{a}^l). \tag{1·1}$$

For the homogeneous deformation the function $\mathbf{r}(\mathbf{a}^l)$ has the form

$$r_i(\mathbf{a}^l) = a_i^l + \sum_{j=1}^{3} s_{ij} a_j^l, \tag{1·2a}$$

the lower indices denoting vector components. s_{ij} is the symmetric strain tensor. For thermal oscillations we have

$$\mathbf{r}(\mathbf{a}^l) = \mathbf{a}^l + \sum_q \mathbf{U}^q e^{i\mathbf{q}\cdot\mathbf{a}^l} \quad (\mathbf{U}^{-q} = \mathbf{U}^{q*}), \tag{1·2b}$$

where the summation extends over all thermal wave vectors $\mathbf{q}$.

We now introduce the conditions of the cyclic lattice. We take a very large piece of the metal, consisting of N lattice points; this will be referred to as the fundamental region. We require that the configuration of the ions and the wave functions of the electrons throughout space be obtained by periodic continuation of the fundamental region.†

Let $2n$ be the number of electrons in the fundamental region, n electrons for each spin direction. We shall denote the n self-consistent wave functions of the electrons by ψ_γ $(\gamma = 1, 2, \ldots, n)$ and normalize them in the fundamental region.

Let $\mathscr{G}(\mathbf{r}, \mathbf{r}') = e^2/|\mathbf{r} - \mathbf{r}'|$ be the interaction between two electrons, and let $V(\mathbf{r}, \mathbf{r}^l) = V(\mathbf{r} - \mathbf{r}^l)$ be the interaction between an electron at $\mathbf{r}$ and an ion at $\mathbf{r}^l$. We introduce the self-consistent potential energy operator V_∞ defined by

$$V_\infty \psi_\gamma = \sum_l^{\infty} V(\mathbf{r}, \mathbf{r}^l)\,\psi_\gamma + 2\sum_{\lambda=1}^{n} (\psi_\lambda \cdot \mathscr{G}\psi_\lambda)_\infty \psi_\gamma - \sum_{\lambda=1}^{n} (\psi_\lambda \cdot \mathscr{G}\psi_\gamma)_\infty \psi_\lambda. \tag{1·3}$$

The suffixes ∞ signify that the lattice summation extends over the whole infinite lattice and that the scalar product is defined as an integral over the whole of space

$$(\psi_\lambda \cdot \mathscr{G}\psi_\gamma)_\infty = \int_\infty d\mathbf{r}'\,\psi_\lambda^*(\mathbf{r}')\,\mathscr{G}(\mathbf{r}, \mathbf{r}')\,\psi_\gamma(\mathbf{r}'). \tag{1·4}$$

Further, we denote the kinetic energy operator of the electron by $T(\mathbf{r})$, i.e.

$$T(\mathbf{r}) = -\frac{\hbar^2}{2m}\sum_{i=1}^{3}\frac{\partial^2}{\partial r_i^2}. \tag{1·5}$$

The Fock-Dirac system of equations for the metallic electrons then reads (cf. I, equations (1) and (2) and § 3)

$$(T + V_\infty - W_\gamma)\,\psi_\gamma = 0 \quad (\gamma = 1, 2, \ldots, n), \tag{1·6}$$

which determines both the wave functions ψ_γ and the parameters W_γ.

Owing to the summation and integration over the whole of space, the operator V_∞ throughout space is obtained by periodic continuation of the fundamental region. This shows that the condition of periodicity imposed upon the wave functions is compatible with the wave equations (1·6).

---

† It can be shown easily that this restriction on the wave functions is justified by applying Bloch's theorem to a lattice with the fundamental region as base.

The first two terms on the right of (1·3) diverge; but their sum, representing the potential of an electrically neutral system, converges. We shall leave the terms in the divergent form, as it is obvious how they should be combined to give a convergent expression.

Let $\phi(\mathbf{r}^l, \mathbf{r}^{l'})$ be the interaction between two ions. The total energy of the fundamental region is the sum of (i) the kinetic energy of the ions, (ii) the kinetic energy of the electrons, (iii) the interactions of the electrons with the ions, (iv) the interactions among the electrons, and (v) the interactions among the ions. After averaging over the rapid motion of the electrons, we can imagine that the metal is replaced by a lattice of ions, with the sum of (i) the average kinetic energy of the electrons, (ii) the average interactions of the electrons with the ions, (iii) the average interactions among the electrons, and (iv) the interactions among the ions as the lattice potential energy. Remembering that we are dealing with $2n$ electrons, n for each spin direction (cf. I, § 3 for the appearance of the factor 2), we have, therefore, the following expression for the potential energy of the fundamental region:

$$\Phi = 2\sum_{\gamma=1}^{n}(\psi_\gamma \,.\, T\psi_\gamma) + 2\sum_{\gamma=1}^{n}\sum_{l}^{\infty}(\psi_\gamma \,.\, V(\mathbf{r},\mathbf{r}^l)\,\psi_\gamma) + 2\sum_{\gamma,\lambda=1}^{n}(\psi_\gamma \,.\, (\psi_\lambda \,.\, \mathscr{G}\psi_\lambda)_\infty\,\psi_\gamma)$$
$$- \sum_{\gamma,\lambda=1}^{n}(\psi_\gamma \,.\, (\psi_\lambda \,.\, \mathscr{G}\psi_\gamma)_\infty\,\psi_\lambda) + \frac{1}{2}\sum_{l}\sum_{l'}^{\infty}\phi(\mathbf{r}^l,\mathbf{r}^{l'}). \qquad (1\cdot7)$$

Here the summation over l without the index ∞ extends over the fundamental region and similarly the scalar products without the index ∞ are defined as integrals over the fundamental region. Referring to the expression (I, equation (5)) for the total energy of the electrons, we see that Φ is simply the sum of the total energy of the electrons in the field of the ions and the interactions among the ions. It depends on the configuration of the ions through the corresponding Fock functions ψ_γ ($\gamma = 1, 2, \ldots, n$) and of course the energy $\phi(\mathbf{r}^l, \mathbf{r}^{l'})$ between two ions; and it varies as the number of atoms in the fundamental region.

2. PERTURBATION METHOD

The perturbation theory for the Fock-Dirac system of equations has been given in I. We shall apply this to develop the lattice potential energy (1·7) with respect to the parameters s_{ij} or $\mathbf{U}^q$ introduced above.

Let ψ_γ denote the solution of the wave equations (1·6) when the ions are displaced in accordance with (1·1), and let $\psi_\gamma^{(0)}$ be the wave functions when the ions are in their equilibrium positions. Now the wave functions oscillate very strongly in the neighbourhood of the ions. When the ions are displaced, they will simply carry the wave functions in their neighbourhood with them. Accordingly we introduce the co-ordinate transformation

$$\mathbf{r} = \mathbf{r}(\boldsymbol{\rho}), \tag{2·1}$$

where the function $\mathbf{r}$ is the same as in (1·1). It follows with (1·1) that the image in $\boldsymbol{\rho}$-space of the position of an ion $\mathbf{r}^l$ is a lattice point $\boldsymbol{\rho}^l = \mathbf{a}^l$. We introduce also the transformed wave functions defined by

$$\overline{\psi}_\gamma(\boldsymbol{\rho}) = [\partial(\mathbf{r})/\partial(\boldsymbol{\rho})]^{\frac{1}{2}}\{\psi_\gamma(\mathbf{r})\}_{\mathbf{r}=\mathbf{r}(\boldsymbol{\rho})}, \tag{2·2}$$

where $\partial(\mathbf{r})/\partial(\boldsymbol{\rho})$ is the Jacobian. As the $\psi_\gamma(\mathbf{r})$ oscillate strongly in the neighbourhood of $\mathbf{r} = \mathbf{r}^l$, so the $\overline{\psi}_\gamma(\boldsymbol{\rho})$ oscillate strongly in the neighbourhood of $\boldsymbol{\rho} = \mathbf{a}^l$. Since the $\psi^{(0)}_\gamma(\mathbf{r})$ oscillate strongly in the neighbourhood of $\mathbf{r} = \mathbf{a}^l$, the above shows that $\psi^{(0)}_\gamma(\boldsymbol{\rho})$ is suitable for a zero approximation to $\overline{\psi}_\gamma(\boldsymbol{\rho})$.

The Jacobian factor is introduced on the right-hand side of (2·2) in order to normalize the $\overline{\psi}_\gamma(\boldsymbol{\rho})$ in the fundamental region in $\boldsymbol{\rho}$-space. This region is obtained from that in $\mathbf{r}$-space by the transformation (2·1) and is the same as the fundamental region of a lattice in the equilibrium configuration.

It follows from (2·1) and (2·2) and the periodic conditions of the wave functions in $\mathbf{r}$-space that the transformed wave functions throughout $\boldsymbol{\rho}$-space are obtained by periodic continuation of the fundamental $\boldsymbol{\rho}$-region.

The transformation (2·2) of the wave functions induces a canonical transformation of operators. We shall write

$$\begin{aligned}\overline{T}(\boldsymbol{\rho})\,\overline{\psi}_\gamma(\boldsymbol{\rho}) &= [\partial(\mathbf{r})/\partial(\boldsymbol{\rho})]^{\frac{1}{2}}\{T(\mathbf{r})\,\psi_\gamma(\mathbf{r})\}_{\mathbf{r}=\mathbf{r}(\boldsymbol{\rho})},\\ \overline{V}(\boldsymbol{\rho},\mathbf{a}^l) &= \{V(\mathbf{r},\mathbf{r}^l)\}_{\mathbf{r}=\mathbf{r}(\boldsymbol{\rho}),\ \mathbf{r}^l=\mathbf{r}(\mathbf{a}^l)},\\ \overline{\mathscr{G}}(\boldsymbol{\rho},\boldsymbol{\rho}') &= \{\mathscr{G}(\mathbf{r},\mathbf{r}')\}_{\mathbf{r}=\mathbf{r}(\boldsymbol{\rho}),\ \mathbf{r}'=\mathbf{r}(\boldsymbol{\rho}')},\\ \overline{\phi}(\mathbf{a}^l,\mathbf{a}^{l'}) &= \{\phi(\mathbf{r}^l,\mathbf{r}^{l'})\}_{\mathbf{r}^l=\mathbf{r}(\mathbf{a}^l),\ \mathbf{r}^{l'}=\mathbf{r}(\mathbf{a}^{l'})},\end{aligned} \tag{2·3}$$

so that the wave equations and the expression for the lattice potential energy are invariant in form. We now abbreviate $\overline{V}(\boldsymbol{\rho},\mathbf{a}^l)$ by $\overline{V}_l(\boldsymbol{\rho})$ and $\overline{\phi}(\mathbf{a}^l,\mathbf{a}^{l'})$ by $\overline{\phi}_{ll'}$; and we expand $\overline{T}(\boldsymbol{\rho})$, $\overline{V}_l(\boldsymbol{\rho})$, $\overline{\mathscr{G}}(\boldsymbol{\rho},\boldsymbol{\rho}')$ and $\overline{\phi}_{ll'}$ into power series of s_{ij} or $\mathbf{U}^q$:

$$\left.\begin{aligned}\overline{T} &= T^{(0)} + T^{(1)} + T^{(2)} + \ldots, & \overline{V}_l &= V_l^{(0)} + V_l^{(1)} + V_l^{(2)} + \ldots,\\ \overline{\mathscr{G}} &= \mathscr{G}^{(0)} + \mathscr{G}^{(1)} + \mathscr{G}^{(2)} + \ldots, & \overline{\phi}_{ll'} &= \phi_{ll'}^{(0)} + \phi_{ll'}^{(1)} + \phi_{ll'}^{(2)} + \ldots.\end{aligned}\right\} \tag{2·4}$$

Evidently we have

$$T^{(0)}(\boldsymbol{\rho}) = T(\boldsymbol{\rho}),\quad V_l^{(0)}(\boldsymbol{\rho}) = V(\boldsymbol{\rho},\mathbf{a}^l),\quad \mathscr{G}^{(0)}(\boldsymbol{\rho},\boldsymbol{\rho}') = \mathscr{G}(\boldsymbol{\rho},\boldsymbol{\rho}'),\quad \phi_{ll'}^{(0)} = \phi(\mathbf{a}^l,\mathbf{a}^{l'}), \tag{2·5}$$

and, corresponding to (1·3) and (1·6), the equations

$$\{T(\boldsymbol{\rho})+V^{(0)}_{\infty}(\boldsymbol{\rho})-W^{(0)}_{\beta}\}\,\psi^{(0)}_{\beta}(\boldsymbol{\rho})=0, \tag{2·6}$$

$$V^{(0)}_{\infty}\psi^{(0)}_{\beta}=\sum_{l}^{\infty}V(\boldsymbol{\rho},\mathbf{a}^{l})\,\psi^{(0)}_{\beta}+2\sum_{\lambda=1}^{n}(\psi^{(0)}_{\lambda}\,.\,\mathscr{G}^{(0)}\psi^{(0)}_{\lambda})_{\infty}\,\psi^{(0)}_{\beta}-\sum_{\lambda=1}^{n}(\psi^{(0)}_{\lambda}\,.\,\mathscr{G}^{(0)}\psi^{(0)}_{\beta})_{\infty}\,\psi^{(0)}_{\lambda}. \tag{2·7}$$

Now since the $\overline{\psi}_{\gamma}(\boldsymbol{\rho})$ and the $\psi^{(0)}_{\gamma}(\boldsymbol{\rho})$ are normalized in the same region and satisfy the same periodic conditions, it is possible to solve the transformed wave equations by expanding the $\overline{\psi}_{\gamma}(\boldsymbol{\rho})$ into series, as in I, of the unperturbed eigenfunctions (I, equations (9) and (11))

$$\overline{\psi}_{\gamma}(\boldsymbol{\rho})=\sum_{\beta}\psi^{(0)}_{\beta}(\boldsymbol{\rho})\,\mathscr{U}_{\beta\gamma},\quad \mathscr{U}_{\beta\gamma}=\delta_{\beta\gamma}+\mathscr{U}^{(1)}_{\beta\gamma}+\mathscr{U}^{(2)}_{\beta\gamma}+\dots. \tag{2·8}$$

The equations for the $\mathscr{U}^{(1)}_{\beta\gamma}$ $(\beta\neq\gamma)$ then read (cf. I, equation (13·1))

$$\begin{aligned}(W^{(0)}_{\beta}-W^{(0)}_{\gamma})\,\mathscr{U}^{(1)}_{\beta\gamma}+\sum_{\lambda=1}^{n}\sum_{\mu\nu}(2\mathscr{G}^{(0)}_{\beta\mu\nu\gamma}-\mathscr{G}^{(0)}_{\beta\mu\gamma\nu})\,(\mathscr{U}^{(1)*}_{\mu\lambda}\delta_{\nu\lambda}+\delta_{\mu\lambda}\mathscr{U}^{(1)}_{\nu\lambda})\\ =-\left\{T^{(1)}_{\beta\gamma}+\sum_{l}^{\infty}V^{(1)}_{l\beta\gamma}+\sum_{\lambda=1}^{n}(2\mathscr{G}^{(1)}_{\beta\lambda\lambda\gamma}-\mathscr{G}^{(1)}_{\beta\lambda\gamma\lambda})\right\}.\end{aligned} \tag{2·9}$$

Here the matrix elements $T^{(i)}_{\beta\gamma}$, $V^{(i)}_{l\beta\gamma}$, $\mathscr{G}^{(i)}_{\beta\mu\nu\gamma}$ $(i=0,1,\dots)$ are defined by

$$\left.\begin{aligned}T^{(i)}_{\beta\gamma}=(\psi^{(0)}_{\beta}\,.\,T^{(i)}\psi^{(0)}_{\gamma}),\qquad V^{(i)}_{l\beta\gamma}=(\psi^{(0)}_{\beta}\,.\,V^{(i)}_{l}\psi^{(0)}_{\gamma}),\\ \mathscr{G}^{(i)}_{\beta\mu\nu\gamma}=(\psi^{(0)}_{\beta}\,.\,(\psi^{(0)}_{\mu}\,.\,\mathscr{G}^{(i)}\psi^{(0)}_{\nu})_{\infty}\,\psi^{(0)}_{\gamma}),\end{aligned}\right\} \tag{2·10}$$

where the scalar products without the index ∞ are volume integrals over the fundamental $\boldsymbol{\rho}$-region, in which the unperturbed eigenfunctions are normalized.

In the calculation of the perturbation energy of the electrons in I, terms have been cancelled by using the relation (I, equation (16))

$$\mathscr{G}^{(i)}_{\beta\mu\nu\gamma}=\mathscr{G}^{(i)}_{\mu\beta\gamma\nu}\quad(i=0,1,\dots),$$

which means, in the present case,

$$(\psi^{(0)}_{\beta}\,.\,(\psi^{(0)}_{\mu}\,.\,\mathscr{G}^{(i)}\psi^{(0)}_{\nu})_{\infty}\,\psi^{(0)}_{\gamma})=(\psi^{(0)}_{\mu}\,.\,(\psi^{(0)}_{\beta}\,.\,\mathscr{G}^{(i)}\psi^{(0)}_{\gamma})_{\infty}\,\psi^{(0)}_{\nu})\quad(i=0,1,\dots). \tag{2·11}$$

This is indeed true, because the $\mathscr{G}^{(i)}(\boldsymbol{\rho},\boldsymbol{\rho}')$ are symmetric in $\boldsymbol{\rho}$ and $\boldsymbol{\rho}'$ and the apparent finite volume integrations implied by the outer brackets of both sides of the above equation, being integrals of periodic quantities over a complete period, can be indefinitely extended. Hence we are permitted to use the final results obtained in I.

For the expansion of the lattice energy $\Phi=\Phi^{(0)}+\Phi^{(1)}+\Phi^{(2)}+\dots$ we have then only to add the expansion of the interactions among the ions to the perturbation expansion of the total energy of the electrons. The first-order and the second-order perturbation energy of the electrons are given by I

(18·1) and (18·3). Since we are dealing with $2n$ electrons, n for each spin direction, these equations are to be slightly modified in accordance with I, § 3. Thus we obtain for $\Phi^{(1)}$, $\Phi^{(2)}$:

$$\Phi^{(1)} = 2\sum_{\gamma=1}^{n} T^{(1)}_{\gamma\gamma} + 2\sum_{\gamma=1}^{n}\sum_{l}^{\infty} V^{(1)}_{l\gamma\gamma} + 2\sum_{\gamma,\lambda=1}^{n} \mathscr{G}^{(1)}_{\gamma\lambda\lambda\gamma} - \sum_{\gamma,\lambda=1}^{n} \mathscr{G}^{(1)}_{\gamma\lambda\gamma\lambda} + \tfrac{1}{2}\sum_{l}\sum_{l'}^{\infty}\phi^{(1)}_{ll'}, \tag{2·12}$$

$$\begin{aligned}\Phi^{(2)} = {} & 2\sum_{\gamma=1}^{n} T^{(2)}_{\gamma\gamma} + 2\sum_{\gamma=1}^{n}\sum_{l}^{\infty} V^{(2)}_{l\gamma\gamma} + 2\sum_{\gamma,\lambda=1}^{n} \mathscr{G}^{(2)}_{\gamma\lambda\lambda\gamma} - \sum_{\gamma,\lambda=1}^{n} \mathscr{G}^{(2)}_{\gamma\lambda\gamma\lambda} + \tfrac{1}{2}\sum_{l}\sum_{l'}^{\infty}\phi^{(2)}_{ll'} \\ & + \text{real part of } 2\sum_{\gamma=1}^{n}\sum_{\beta}\left\{T'^{(1)}_{\gamma\beta} + \sum_{l}^{\infty} V^{(1)}_{l\gamma\beta} + 2\sum_{\lambda=1}^{n}\mathscr{G}^{(1)}_{\gamma\lambda\lambda\beta} - \sum_{\lambda=1}^{n}\mathscr{G}^{(1)}_{\gamma\lambda\beta\lambda}\right\}\mathscr{U}^{(1)}_{\beta\gamma}.\end{aligned} \tag{2·13}$$

The potential energy $\Phi^{(0)}$ of the lattice in equilibrium can be obtained from (1·7) by replacing $\mathbf{r}^l$ and ψ_γ by $\mathbf{a}^l$ and $\psi^{(0)}_\gamma$.

3. THE SUBSTITUTION OPERATOR

The explicit expressions for $V^{(i)}_l$, $\mathscr{G}^{(i)}$, $\phi^{(i)}_{ll'}$ ($i = 1, 2, \ldots$) are easily obtained from the Taylor expansions of $\overline{V}_l$, $\overline{\mathscr{G}}$, $\overline{\phi}_{ll'}$ of (2·3). Remembering that the interactions V, $\mathscr{G}$, ϕ depend only on the relative co-ordinates of the interacting particles, we have for $V^{(1)}_l$, $V^{(2)}_l$

$$V^{(1)}_l(\boldsymbol{\rho}) = \sum_{i=1}^{3}\{r_i(\boldsymbol{\rho}) - r_i(\mathbf{a}^l) - \rho_i + a^l_i\}\frac{\partial V(\boldsymbol{\rho}-\mathbf{a}^l)}{\partial(\rho_i - a^l_i)}, \tag{3·1}$$

$$V^{(2)}_l(\boldsymbol{\rho}) = \frac{1}{2!}\sum_{i,j=1}^{3}\{r_i(\boldsymbol{\rho}) - r_i(\mathbf{a}^l) - \rho_i + a^l_i\}\{r_j(\boldsymbol{\rho}) - r_j(\mathbf{a}^l) - \rho_j + a^l_j\}\frac{\partial^2 V(\boldsymbol{\rho}-\mathbf{a}^l)}{\partial(\rho_i - a^l_i)\,\partial(\rho_j - a^l_j)}, \tag{3·2}$$

and similarly for $\mathscr{G}^{(1)}$, $\mathscr{G}^{(2)}$, $\phi^{(1)}_{ll'}$, $\phi^{(2)}_{ll'}$.

The explicit expressions for $T^{(1)}$ and $T^{(2)}$ are to be calculated from the expansion of $\overline{T}$, which, by (2·2) and (2·3), is given by

$$\left.\begin{aligned}\overline{T}(\boldsymbol{\rho}) &= \left(\frac{\partial(\mathbf{r})}{\partial(\boldsymbol{\rho})}\right)^{\frac{1}{2}}\{T(\mathbf{r})\}_{\mathbf{r}=\mathbf{r}(\boldsymbol{\rho})}\left\{\frac{\partial(\mathbf{r})}{\partial(\boldsymbol{\rho})}\right\}^{-\frac{1}{2}}, \\ \{T(\mathbf{r})\}_{\mathbf{r}=\mathbf{r}(\boldsymbol{\rho})} &= -\frac{\hbar^2}{2m}\sum_{h,i,j=1}^{3}\left(\frac{\partial\rho_i}{\partial r_h}\right)\frac{\partial}{\partial\rho_i}\left(\frac{\partial\rho_j}{\partial r_h}\right)\frac{\partial}{\partial\rho_j},\end{aligned}\right\} \tag{3·3}$$

where the $(\partial\rho_i/\partial r_h)$ are expressed in terms of $\boldsymbol{\rho}$.

We shall give another method for the calculation of $V^{(i)}_l$ and $T^{(i)}$ ($i = 1, 2, \ldots$) which yields at the same time certain useful relations. We introduce the substitution operator

$$\begin{aligned}D_l(\rho) = c\left(\frac{\partial(\mathbf{r})}{\partial(\boldsymbol{\rho})}\right)^{\frac{1}{2}}\Bigg[& 1 + \sum_{i=1}^{3}\{r_i(\boldsymbol{\rho}) - r_i(\mathbf{a}^l) - \rho_i + a^l_i\}\frac{\partial}{\partial(\rho_i - a^l_i)} \\ & + \frac{1}{2!}\sum_{i,j=1}^{3}\{r_i(\boldsymbol{\rho}) - r_i(\mathbf{a}^l) - \rho_i + a^l_i\}\{r_j(\boldsymbol{\rho}) - r_j(\mathbf{a}^l) - \rho_j + a^l_j\}\frac{\partial^2}{\partial(\rho_i - a^l_i)\,\partial(\rho_j - a^l_j)} + \ldots\Bigg],\end{aligned} \tag{3·4}$$

c being an arbitrary constant to be chosen later. For an arbitrary function $f(\boldsymbol{\rho}-\mathbf{a}^l)$, we have then

$$D_l(\boldsymbol{\rho})\,f(\boldsymbol{\rho}-\mathbf{a}^l) = c[\partial(\mathbf{r})/\partial(\boldsymbol{\rho})]^{\frac{1}{2}}\,\{f(\mathbf{r}-\mathbf{r}^l)\}_{\mathbf{r}=\mathbf{r}(\boldsymbol{\rho}),\ \mathbf{r}^l=\mathbf{r}(\mathbf{a}^l)}. \tag{3·5}$$

Now both $T(\boldsymbol{\rho})\,f(\boldsymbol{\rho}-\mathbf{a}^l)$ and $V(\boldsymbol{\rho}-\mathbf{a}^l)\,f(\boldsymbol{\rho}-\mathbf{a}^l)$ are functions of $\boldsymbol{\rho}-\mathbf{a}^l$; we have therefore also

$$D_l(\boldsymbol{\rho})\,T(\boldsymbol{\rho})\,f(\boldsymbol{\rho}-\mathbf{a}^l) = c[\partial(\mathbf{r})/\partial(\boldsymbol{\rho})]^{\frac{1}{2}}\,\{T(\mathbf{r})\,f(\mathbf{r}-\mathbf{r}^l)\}_{\mathbf{r}=\mathbf{r}(\boldsymbol{\rho}),\ \mathbf{r}^l=\mathbf{r}(\mathbf{a}^l)}, \tag{3·6}$$

and

$$D_l(\boldsymbol{\rho})\,V(\boldsymbol{\rho}-\mathbf{a}^l)\,f(\boldsymbol{\rho}-\mathbf{a}^l) = c[\partial(\mathbf{r})/\partial(\boldsymbol{\rho})]^{\frac{1}{2}}\,\{V(\mathbf{r}-\mathbf{r}^l)\,f(\mathbf{r}-\mathbf{r}^l)\}_{\mathbf{r}=\mathbf{r}(\boldsymbol{\rho}),\ \mathbf{r}^l=\mathbf{r}(\mathbf{a}^l)}. \tag{3·7}$$

The right-hand sides of these equations are, by (3·5) and (2·3), equal to $\overline{T}(\boldsymbol{\rho})\,D_l(\boldsymbol{\rho})\,f(\boldsymbol{\rho}-\mathbf{a}^l)$ and $\overline{V}_l(\boldsymbol{\rho})\,D_l(\boldsymbol{\rho})\,f(\boldsymbol{\rho}-\mathbf{a}^l)$ respectively. Hence, omitting the arbitrary function $f(\boldsymbol{\rho}-\mathbf{a}^l)$, we obtain the operator equations

$$D_l(\boldsymbol{\rho})\,T(\boldsymbol{\rho}) = \overline{T}(\boldsymbol{\rho})\,D_l(\boldsymbol{\rho}), \quad D_l(\boldsymbol{\rho})\,V(\boldsymbol{\rho}-\mathbf{a}^l) = \overline{V}_l(\boldsymbol{\rho})\,D_l(\boldsymbol{\rho}). \tag{3·8}$$

From these, with the series expansions of $\overline{T}$, $\overline{V}_l$, and D_l (the constant c of (3·4) being chosen so that $D_l = 1 + D_l^{(1)} + D_l^{(2)} + \ldots$), follow

$$T^{(1)} = D_l^{(1)}T - TD_l^{(1)}, \quad T^{(2)} + T^{(1)}D_l^{(1)} = D_l^{(2)}T - TD_l^{(2)}, \tag{3·9}$$

and similar relations for expressing $V_l^{(1)}$, $V_l^{(2)}$ in terms of $V(\boldsymbol{\rho}-\mathbf{a}^l)$ and $D_l^{(1)}$, $D_l^{(2)}$. These equations can be used for the calculations of $T^{(1)}$, $T^{(2)}$ and $V_l^{(1)}$, $V_l^{(2)}$.

When l_1, l_2, l_3 all assume the value zero, we shall omit the index l of D_l or V_l. The lattice point a^l coincides in this case with the origin, so that $V(\boldsymbol{\rho}-\mathbf{a}^l)$ becomes simply $V(\boldsymbol{\rho})$. Adding to (3·9) the corresponding equations with V in place of T and choosing in particular $l_1 = l_2 = l_3 = 0$, we get

$$T^{(1)} + V^{(1)} = D^{(1)}(T+V) - (T+V)\,D^{(1)}, \tag{3·10}$$

$$T^{(2)} + V^{(2)} + (T^{(1)} + V^{(1)})\,D^{(1)} = D^{(2)}(T+V) - (T+V)\,D^{(2)}. \tag{3·11}$$

These† will be of use later for the transformation of certain volume integrals.

4. THE PERTURBATION EQUATIONS FOR HOMOGENEOUS DEFORMATION

From now on we shall confine ourselves to the case of homogeneous deformation. The co-ordinate transformation $\mathbf{r}(\boldsymbol{\rho})$ is a homogeneous linear transformation with constant coefficients

$$r_i(\boldsymbol{\rho}) = \rho_i + \sum_{j=1}^{3} s_{ij}\rho_j. \tag{4·1}$$

---

† Alternatively, these relations can be derived by applying the perturbation method of § 2 to the problem of an isolated atom situated at the origin.

The quantities $V_l^{(i)}(\boldsymbol{\rho})$, $\mathscr{G}^{(i)}(\boldsymbol{\rho}, \boldsymbol{\rho}')$, $\phi_{ll'}^{(i)}$ $(i = 0, 1, 2, \ldots)$ depend only on the relative co-ordinates $\boldsymbol{\rho} - \mathbf{a}^l$, $\boldsymbol{\rho} - \boldsymbol{\rho}'$, $\mathbf{a}^l - \mathbf{a}^{l'}$ respectively. For instance, we have

$$V_l^{(1)} = \sum_{i,j=1}^{3} s_{ij}(\rho_j - a_j^l) \frac{\partial V(\boldsymbol{\rho} - \mathbf{a}^l)}{\partial(\rho_i - a_i^l)}, \tag{4·2}$$

etc.

The partial derivatives $\partial r_i/\partial \rho_j$ and $\partial \rho_i/\partial r_j$ are constants. Hence the Jacobian factor is a constant and, by (3·3), the transformed kinetic energy operator is a homogeneous quadratic form of $\partial/\partial \rho_i$ $(i = 1, 2, 3)$ with constant coefficients, viz.

$$\overline{T} = -\frac{\hbar^2}{2m} \left\{ \sum_{i=1}^{3} \frac{\partial^2}{\partial \rho_i^2} - 2 \sum_{i,j=1}^{3} s_{ij} \frac{\partial^2}{\partial \rho_i \partial \rho_j} + 3 \sum_{i,j,q=1}^{3} s_{ij} s_{jq} \frac{\partial^2}{\partial \rho_i \partial \rho_q} - \ldots \right\}. \tag{4·3}$$

The constant c occurring in (3·7) may be chosen in such a way to cancel the Jacobian factor occurring also there. Then, choosing $l_1 = l_2 = l_3 = 0$, we have

$$D^{(1)} = \sum_{i,j=1}^{3} s_{ij} \rho_j \frac{\partial}{\partial \rho_i}, \quad D^{(2)} = \frac{1}{2!} \sum_{i,j,p,q=1}^{3} s_{ij} s_{pq} \rho_j \rho_q \frac{\partial^2}{\partial \rho_i \partial \rho_p}. \tag{4·4}$$

The unperturbed eigenfunctions are known to be of the form (Bloch's theorem (1928))

$$\psi_\beta^{(0)}(\boldsymbol{\rho}) = N^{-\frac{1}{2}} e^{i\mathbf{k}.\boldsymbol{\rho}} u_\beta^{(0)}(\boldsymbol{\rho}), \tag{4·5}$$

where the $u_\beta^{(0)}$ possess the periods $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, the lattice vectors of the undeformed lattice, and the $\mathbf{k}$ are the corresponding reduced wave vectors. The vectors $\mathbf{k}$ assume practically continuous values; therefore the eigenvalues $W_\beta^{(0)}$ are divided into bands. Each band contains N levels; each level belongs to a different wave vector. Belonging to one and the same wave vector there are an infinite number of eigenfunctions, one from each band, their eigenvalues being separated considerably from each other. Occasionally we shall write $\beta = (b, \mathbf{k})$ with b and $\mathbf{k}$ indicating respectively the band and the reduced wave vector.

Now $\partial^2 \psi_\beta^{(0)}/\partial \rho_i \partial \rho_j$ has the same wave character as $\psi_\beta^{(0)}$ itself; and the eigenfunctions belonging to different wave vectors are orthogonal. Hence $T_{\beta\gamma}^{(i)}$ vanishes except when $\psi_\beta^{(0)}$ belongs to the same wave vector as $\psi_\gamma^{(0)}$. And the same is true for $\mathscr{G}_{\beta\lambda\lambda\gamma}^{(i)}$, $\mathscr{G}_{\beta\lambda\gamma\lambda}^{(i)}$ and the infinite summation over l of $V_{l\beta\gamma}^{(i)}$, remembering that $\mathscr{G}^{(i)}$ and $V_l^{(i)}$ depend respectively only on the relative co-ordinates $\boldsymbol{\rho} - \boldsymbol{\rho}'$ and $\boldsymbol{\rho} - \mathbf{a}^l$. That is to say that we have

$$T_{\beta\gamma}^{(i)} = 0, \quad \sum_{l}^{\infty} V_{l\beta\gamma}^{(i)} = 0, \quad \mathscr{G}_{\beta\lambda\lambda\gamma}^{(i)} = 0, \quad \mathscr{G}_{\beta\lambda\gamma\lambda}^{(i)} = 0 \quad (i = 0, 1, \ldots), \tag{4·6}$$

unless

$$\beta = (b, \mathbf{k}), \quad \gamma = (b', \mathbf{k}') \quad \text{and} \quad \mathbf{k} = \mathbf{k}'. \tag{4·7}$$

Owing to these facts it follows from the equation (2·9) for $\mathscr{U}^{(1)}_{\beta\gamma}$ and the similar equations for $\mathscr{U}^{(2)}_{\beta\gamma}$, etc., that the $\mathscr{U}^{(i)}_{\beta\gamma}$ $(i = 1, 2, \ldots)$ vanish unless the condition (4·7) holds. Thus the transformed wave functions $\bar{\psi}_\beta(\boldsymbol{\rho})$ for the electrons in the deformed lattice are again of the form (4·5) with some periodic functions $\bar{u}_\beta$, say, in place of the $u^{(0)}_\beta$. This is equivalent to the Bloch theorem for the deformed lattice which asserts that the wave functions $\psi_\beta(\mathbf{r})$ are of the form

$$\psi_\beta(\mathbf{r}) = N^{-\frac{1}{2}} e^{i\bar{\mathbf{k}}.\mathbf{r}} u_\beta(\mathbf{r}), \tag{4·8}$$

where $\bar{\mathbf{k}}$ denotes the wave vectors of the deformed lattice and $u_\beta(\mathbf{r})$ possesses the periods $\bar{\mathbf{a}}_1$, $\bar{\mathbf{a}}_2$, $\bar{\mathbf{a}}_3$, the lattice vectors of the deformed lattice.

Therefore we have to calculate those matrix elements for which the wave vectors satisfy the condition (4·7). All equations below in which both β and γ occur refer only to such values of β, γ as satisfy (4·7).

We shall subdivide the fundamental $\boldsymbol{\rho}$-region into N identical polyhedra, the lth one surrounding the lth ion. Since the integrand of $T^{(i)}_{\beta\gamma}$ possesses the periods $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, the integral over the fundamental $\boldsymbol{\rho}$-region is simply N times the integral over the region of one atomic polyhedron. We have therefore

$$T^{(i)}_{\beta\gamma} = N\int_\Omega d\boldsymbol{\rho}\, \psi^{(0)*}_\beta(\boldsymbol{\rho})\, T^{(i)}(\boldsymbol{\rho})\, \psi^{(0)}_\gamma(\boldsymbol{\rho}), \tag{4·9}$$

where the integral with the index Ω covers the region of an atomic polyhedron, say, the basic one that surrounds the ion at the origin ($\mathbf{a}^l = 0$). Similarly, we have

$$\begin{aligned}\sum_l^\infty V^{(i)}_{l\beta\gamma} &= N\sum_l^\infty \int_\Omega d\boldsymbol{\rho}\, \psi^{(0)*}_\beta V^{(i)}_l \psi^{(0)}_\gamma \\ &= N\int_\Omega d\boldsymbol{\rho}\, \psi^{(0)*}_\beta V^{(i)} \psi^{(0)}_\gamma + N\sum_l^{\infty}{}' \int_\Omega d\boldsymbol{\rho}\, \psi^{(0)*}_\beta V^{(i)}_l \psi^{(0)}_\gamma,\end{aligned} \tag{4·10}$$

where the accent to the summation sign indicates that the term with $l_1 = l_2 = l_3 = 0$ is to be omitted in the lattice summation.

Making use of these relations, we collect here the perturbation equations for $\mathscr{U}^{(1)}_{\beta\gamma}$, $\Phi^{(1)}$, $\Phi^{(2)}$ (cf. (2·9), (2·12), (2·13) respectively)

$$\begin{aligned}&(W^{(0)}_\beta - W^{(0)}_\gamma)\,\mathscr{U}^{(1)}_{\beta\gamma} + \sum_{\lambda=1}^n \sum_{\mu\nu} (2\mathscr{G}^{(0)}_{\beta\mu\nu\gamma} - \mathscr{G}^{(0)}_{\beta\mu\gamma\nu})\,(\mathscr{U}^{(1)*}_{\mu\lambda}\delta_{\nu\lambda} + \delta_{\mu\lambda}\mathscr{U}^{(1)}_{\nu\lambda}) \\ &\quad = -N\int_\Omega d\boldsymbol{\rho}\, \psi^{(0)*}_\beta (T^{(1)} + V^{(1)})\, \psi^{(0)}_\gamma \\ &\qquad -\left\{\sum_l^{\infty}{}' N\int_\Omega d\boldsymbol{\rho}\, \psi^{(0)*}_\beta V^{(1)}_l \psi^{(0)}_\gamma + \sum_{\lambda=1}^n (2\mathscr{G}^{(1)}_{\beta\lambda\lambda\gamma} - \mathscr{G}^{(1)}_{\beta\lambda\gamma\lambda})\right\} \quad (\beta \neq \gamma),\end{aligned} \tag{4·11}$$

$$\begin{aligned}\Phi^{(1)} = {} & 2\sum_{\gamma=1}^{n} N\int_{\Omega} d\boldsymbol{\rho}\, \psi_{\gamma}^{(0)*}(T^{(1)}+V^{(1)})\,\psi_{\gamma}^{(0)} \\ & + 2\sum_{\gamma=1}^{n}\sum_{l}^{\infty}{}' N\int_{\Omega} d\boldsymbol{\rho}\, \psi_{\gamma}^{(0)*} V_{l}^{(1)}\psi_{\gamma}^{(0)} + 2\sum_{\gamma,\lambda=1}^{n}\mathscr{G}^{(1)}_{\gamma\lambda\lambda\gamma} - \sum_{\gamma,\lambda=1}^{n}\mathscr{G}^{(1)}_{\gamma\lambda\gamma\lambda} + \tfrac{1}{2}\sum_{l}\sum_{l'}^{\infty}\phi^{(1)}_{ll'},\end{aligned} \tag{4·12}$$

$$\begin{aligned}\Phi^{(2)} = {} & 2\sum_{\gamma=1}^{n} N\int_{\Omega} d\boldsymbol{\rho}\, \psi_{\gamma}^{(0)*}(T^{(2)}+V^{(2)})\,\psi_{\gamma}^{(0)} \\ & + 2\sum_{\gamma=1}^{n}\sum_{l}^{\infty}{}' N\int_{\Omega} d\boldsymbol{\rho}\, \psi_{\gamma}^{(0)*} V_{l}^{(2)}\psi_{\gamma}^{(0)} + 2\sum_{\gamma,\lambda=1}^{n}\mathscr{G}^{(2)}_{\gamma\lambda\lambda\gamma} - \sum_{\gamma,\lambda=1}^{n}\mathscr{G}^{(2)}_{\gamma\lambda\gamma\lambda} + \tfrac{1}{2}\sum_{l}\sum_{l'}^{\infty}\phi^{(2)}_{ll'} \\ & + \text{real part of } 2\sum_{\gamma=1}^{n}\sum_{\beta}\left\{T^{(1)}_{\gamma\beta} + \sum_{l}^{\infty} V^{(1)}_{l\gamma\beta} + 2\sum_{\lambda=1}^{n}\mathscr{G}^{(1)}_{\gamma\lambda\lambda\beta} - \sum_{\lambda=1}^{n}\mathscr{G}^{(1)}_{\gamma\lambda\beta\lambda}\right\}\mathscr{U}^{(1)}_{\beta\gamma}.\end{aligned} \tag{4·13}$$

The unperturbed eigenvalue problem (2·6) is more conveniently written in this case

$$(T+V+V_{\infty}^{(0)\prime}-W_{\beta}^{(0)})\,\psi_{\beta}^{(0)} = 0, \tag{4·14}$$

where the accent to $V_{\infty}^{(0)}$ indicates that the term representing the potential due to the ion at $\mathbf{a}^{l} = 0$ is missing from the summation of (2·7).

5. Solution of the perturbation equations

For the solution of the perturbation equations we shall use the following method, which is in fact identical with the method of operator calculus introduced by one of us (Fuchs 1940); but for the present purpose it is more convenient to avoid the abstract notation used in that paper.

We consider an integral over a bounded region Ω of the form

$$\int_{\Omega} d\boldsymbol{\rho}(T^{(i)}f^{*}).g \quad (i = 0, 1, \ldots),$$

where f and g are arbitrary functions of $\boldsymbol{\rho}$. As $T^{(i)}$ is, by (4·3), a certain linear combination of $\partial^2/\partial\rho_p\partial\rho_q$ $(p, q = 1, 2, 3)$, we can transform the above integral, by two partial integrations, into an integral with the integrand $f^{*}(T^{(i)}g)$ plus certain surface integrals. This transformation will often be needed in the following (in the case where $i = 0$, this coincides with Green's theorem):

(*a*) *The equation for* $\mathscr{U}^{(1)}_{\beta\gamma}$. From (4·14) follows

$$\begin{aligned}(W_{\beta}^{(0)} - W_{\gamma}^{(0)})\int_{\Omega} d\boldsymbol{\rho}\, \psi_{\beta}^{(0)*}\,.\,D^{(i)}\psi_{\gamma}^{(0)} \\ = \int_{\Omega} d\boldsymbol{\rho}(T+V+V_{\infty}^{(0)\prime})\,\psi_{\beta}^{(0)*}\,.\,D^{(i)}\psi_{\gamma}^{(0)} - \int_{\Omega} d\boldsymbol{\rho}\,\psi_{\beta}^{(0)*}\,.\,D^{(i)}(T+V+V_{\infty}^{(0)\prime})\,\psi_{\gamma}^{(0)}.\end{aligned}$$

By partial integrations as explained above, we transform the integral with the integrand $T\psi_\beta^{(0)*}.D^{(i)}\psi_\gamma^{(0)}$ into one with the integrand $\psi_\beta^{(0)*}.TD^{(i)}\psi_\gamma^{(0)}$ and thus get

$$\begin{aligned}(W_\beta^{(0)}-W_\gamma^{(0)})\int_\Omega d\boldsymbol{\rho}\,\psi_\beta^{(0)*}.D^{(i)}\psi_\gamma^{(0)} &= \int_\Omega d\boldsymbol{\rho}(V_\infty^{(0)\prime}\psi_\beta^{(0)*}.D^{(i)}\psi_\gamma^{(0)}-\psi_\beta^{(0)*}.D^{(i)}V_\infty^{(0)\prime}\psi_\gamma^{(0)})\\ &\quad-\int_\Omega d\boldsymbol{\rho}\,\psi_\beta^{(0)*}.[D^{(i)}(T+V)-(T+V)D^{(i)}]\,\psi_\gamma^{(0)}\\ &\quad+\frac{\hbar^2}{2m}\sum_{j=1}^{3}\int_\Omega d\sigma\,n_j\left(\psi_\beta^{(0)*}.\frac{\partial D^{(i)}\psi_\gamma^{(0)}}{\partial\rho_j}-\frac{\partial\psi_\beta^{(0)*}}{\partial\rho_j}D^{(i)}\psi_\gamma^{(0)}\right).\end{aligned}\tag{5·1}$$

Here the integral over $d\sigma$ denotes integration over the surface of the basic polyhedron, and the n_j denote the components of the unit vector in the direction of the outward normal to this surface.

Now in case that $i=1$, the operator within the square bracket of (5·1) is, by (3·10), equal to $(T^{(1)}+V^{(1)})$. Hence we can use (5·1) to transform the integral over Ω of $\psi_\beta^{(0)*}.(T^{(1)}+V^{(1)})\,\psi_\gamma^{(0)}$ and thus obtain from (4·11) the following equation determining $\mathscr{U}_{\beta\gamma}^{(1)}$ $(\beta\neq\gamma)$:

$$\begin{aligned}(W_\beta^{(0)}-W_\gamma^{(0)})\,\mathscr{U}_{\beta\gamma}^{(1)}+\sum_{\lambda=1}^{n}\sum_{\mu\nu}(2\mathscr{G}_{\beta\mu\nu\gamma}^{(0)}-\mathscr{G}_{\beta\mu\gamma\nu}^{(0)})(\mathscr{U}_{\mu\lambda}^{(1)*}\delta_{\nu\lambda}+\delta_{\mu\lambda}\mathscr{U}_{\nu\lambda}^{(1)})\\ =(W_\beta^{(0)}-W_\gamma^{(0)})\,N\int_\Omega d\boldsymbol{\rho}\,\psi_\beta^{(0)*}.D^{(1)}\psi_\gamma^{(0)}\\ +N\int_\Omega d\boldsymbol{\rho}(\psi_\beta^{(0)*}.D^{(1)}V_\infty^{(0)\prime}\psi_\gamma^{(0)}-V_\infty^{(0)\prime}\psi_\beta^{(0)*}.D^{(1)}\psi_\gamma^{(0)})\\ +\frac{\hbar^2}{2m}\sum_{i=1}^{3}N\int_\Omega d\sigma\,n_i\left(\frac{\partial\psi_\beta^{(0)*}}{\partial\rho_i}.D^{(1)}\psi_\gamma^{(0)}-\psi_\beta^{(0)*}.\frac{\partial D^{(1)}\psi_\gamma^{(0)}}{\partial\rho_i}\right)\\ -\left\{\sum_l^{\infty}{}'\,N\int_\Omega d\boldsymbol{\rho}\,\psi_\beta^{(0)*}V_l^{(1)}\psi_\gamma^{(0)}+\sum_{\lambda=1}^{n}(2\mathscr{G}_{\beta\lambda\lambda\gamma}^{(1)}-\mathscr{G}_{\beta\lambda\gamma\lambda}^{(1)})\right\}.\end{aligned}\tag{5·2}$$

For the calculation of $\Phi^{(2)}$ we need only such $\mathscr{U}_{\beta\gamma}^{(1)}$ that $\psi_\gamma^{(0)}$ refers to an occupied state. Employing plane-wave eigenfunctions according to the approximation of the metallic correspondence (Slater 1934, § 18), we have estimated that the second term on the left of (5·2) is in general much smaller than the first, so that (5·2) can be solved by the method of successive approximations. From the zero approximation to $\mathscr{U}_{\beta\gamma}^{(1)}$ $(\beta\neq\gamma)$,

$$\begin{aligned}\mathscr{U}_{\beta\gamma}^{(1)} &= N\int_\Omega d\boldsymbol{\rho}\,\psi_\beta^{(0)*}.D^{(1)}\psi_\gamma^{(0)}\\ &\quad+\frac{1}{W_\beta^{(0)}-W_\gamma^{(0)}}\left\{\frac{\hbar^2}{2m}\sum_{i=1}^{3}N\int_\Omega d\sigma\,n_i\left(\frac{\partial\psi_\beta^{(0)*}}{\partial\rho_i}.D^{(1)}\psi_\gamma^{(0)}-\psi_\beta^{(0)*}.\frac{\partial D^{(1)}\psi_\gamma^{(0)}}{\partial\rho_i}\right)\right.\\ &\quad+N\int_\Omega d\boldsymbol{\rho}(\psi_\beta^{(0)*}.D^{(1)}V_\infty^{(0)\prime}\psi_\gamma^{(0)}-V_\infty^{(0)\prime}\psi_\beta^{(0)*}.D^{(1)}\psi_\gamma^{(0)})\\ &\quad\left.-\sum_l^{\infty}{}'\,N\int_\Omega d\boldsymbol{\rho}\,\psi_\beta^{(0)*}.V_l^{(1)}\psi_\gamma^{(0)}-\sum_{\lambda=1}^{n}(2\mathscr{G}_{\beta\lambda\lambda\gamma}^{(1)}-\mathscr{G}_{\beta\lambda\gamma\lambda}^{(1)})\right\},\end{aligned}\tag{5·3}$$

we can proceed to the first approximation, and so on. Hence if we write

$$\mathscr{U}^{(1)}_{\beta\gamma} = N\int_{\Omega} d\boldsymbol{\rho}\, \psi^{(0)*}_{\beta} . D^{(1)} \psi^{(0)}_{\gamma} + X^{(1)}_{\beta\gamma}, \tag{5·4}$$

then $X^{(1)}_{\beta\gamma}$ $(\beta \neq \gamma)$, which can be obtained by successive approximations, is in general small on account of the big denominator $W^{(0)}_{\beta} - W^{(0)}_{\gamma}$.

For $\beta = \gamma$, from the normalization conditions follows $\mathscr{U}^{(1)}_{\gamma\gamma} = 0$ (cf. I, 14)).

(*b*) *The first-order lattice energy.* The first-order lattice energy is given by (4·12), i.e.

$$\begin{aligned}\Phi^{(1)} = 2\sum_{\gamma=1}^{n} E^{(1)}_{\gamma} + 2\sum_{\gamma=1}^{n}\sum_{l}^{\infty}{}' N\int_{\Omega} d\boldsymbol{\rho}\, \psi^{(0)*}_{\gamma} . V^{(1)}_{l} \psi^{(0)}_{\gamma} \\ + 2\sum_{\gamma,\lambda=1}^{n} \mathscr{G}^{(1)}_{\gamma\lambda\lambda\gamma} - \sum_{\gamma,\lambda=1}^{n} \mathscr{G}^{(1)}_{\gamma\lambda\gamma\lambda} + \tfrac{1}{2}\sum_{l}\sum_{l'}^{\infty} \phi^{(1)}_{ll'},\end{aligned} \tag{5·5}$$

where $E^{(1)}_{\gamma}$ stands for the integral of $N\psi^{(0)*}_{\gamma} . (T^{(1)} + V^{(1)})\, \psi^{(0)}_{\gamma}$. Putting $\beta = \gamma$ and $i = 1$ in (5·1) and remembering (3·10), we have for $E^{(1)}_{\gamma}$

$$\begin{aligned}E^{(1)}_{\gamma} = N\int_{\Omega} d\boldsymbol{\rho} (V^{(0)\prime}_{\infty} \psi^{(0)*}_{\gamma} . D^{(1)} \psi^{(0)}_{\gamma} - \psi^{(0)*}_{\gamma} . D^{(1)} V^{(0)\prime}_{\infty} \psi^{(0)}_{\gamma}) \\ + \frac{\hbar^2}{2m}\sum_{i=1}^{3}\int_{\Omega} d\sigma\, n_i \left(\psi^{(0)*}_{\gamma} . \frac{\partial D^{(1)} \psi^{(0)}_{\gamma}}{\partial \rho_i} - \frac{\partial \psi^{(0)*}_{\gamma}}{\partial \rho_i} D^{(1)} \psi^{(0)}_{\gamma}\right).\end{aligned} \tag{5·6}$$

(*c*) *The second-order lattice energy.* The second-order lattice energy is given by (4·13). We introduce the operator

$$H^{(1)} = T^{(1)} + \sum_{l}^{\infty} V^{(1)}_{l} + 2\sum_{\lambda=1}^{n} \mathscr{G}^{(1)}_{\lambda\lambda}, \tag{5·7}$$

where $\mathscr{G}^{(1)}_{\lambda\lambda}$ is defined by

$$\mathscr{G}^{(1)}_{\lambda\lambda}(\boldsymbol{\rho}) = \int_{\infty} d\boldsymbol{\rho}'\, \psi^{(0)*}_{\lambda}(\boldsymbol{\rho}')\, \mathscr{G}^{(1)}(\boldsymbol{\rho}, \boldsymbol{\rho}')\, \psi^{(0)}_{\lambda}(\boldsymbol{\rho}'). \tag{5·8}$$

We can then write (4·13) in the form

$$\Phi^{(2)} = \Phi^{(2)\prime} + \text{real part of } 2\sum_{\gamma=1}^{n}\left\{N\int_{\Omega} d\boldsymbol{\rho}\, \psi^{(0)*}_{\gamma} . (T^{(2)} + V^{(2)})\, \psi^{(0)}_{\gamma} + \sum_{\beta} H^{(1)}_{\gamma\beta} \mathscr{U}^{(1)}_{\beta\gamma}\right\}, \tag{5·9}$$

where $\Phi^{(2)\prime}$ is but an abbreviation for those terms of (4·13) which will not be considered for the moment, i.e.

$$\begin{aligned}\Phi^{(2)\prime} = 2\sum_{\gamma=1}^{n}\sum_{l}^{\infty}{}' N\int_{\Omega} d\boldsymbol{\rho}\, \psi^{(0)*}_{\gamma} V^{(2)}_{l} \psi^{(0)}_{\gamma} + 2\sum_{\gamma,\lambda=1}^{n} \mathscr{G}^{(2)}_{\gamma\lambda\lambda\gamma} - \sum_{\gamma,\lambda=1}^{n} \mathscr{G}^{(2)}_{\gamma\lambda\gamma\lambda} + \tfrac{1}{2}\sum_{l}\sum_{l'}^{\infty} \phi^{(2)}_{ll'} \\ - \text{real part of } 2\sum_{\gamma,\lambda=1}^{n}\sum_{\beta} \mathscr{G}^{(1)}_{\gamma\lambda\beta\lambda} \mathscr{U}^{(1)}_{\beta\gamma}.\end{aligned} \tag{5·10}$$

With $\mathscr{U}^{(1)}_{\beta\gamma}$ given by (5·4), we have to consider the sum

$$\sum_{\beta} H^{(1)}_{\gamma\beta}\mathscr{U}^{(1)}_{\beta\gamma} = N\sum_{\beta\neq\gamma} H^{(1)}_{\gamma\beta}\int_{\Omega} d\boldsymbol{\rho}\,\psi^{(0)*}_{\beta}.\,D^{(1)}\psi^{(0)}_{\gamma} + \sum_{\beta\neq\gamma} H^{(1)}_{\gamma\beta}X^{(1)}_{\beta\gamma} \tag{5·11}$$

of (5·9). Now $H^{(1)}$ is self-adjoint. Furthermore, $H^{(1)}\psi^{(0)*}_{\gamma}$ has the same wave character as $\psi^{(0)*}_{\gamma}$ itself. Therefore $H^{(1)}\psi^{(0)*}_{\gamma}$ can be expanded as follows:

$$H^{(1)}\psi^{(0)*}_{\gamma} = \sum_{\beta}(H^{(1)}\psi^{(0)}_{\gamma}.\,\psi^{(0)}_{\beta})\,\psi^{(0)*}_{\beta} = \sum_{\beta}(\psi^{(0)}_{\gamma}.\,H^{(1)}\psi^{(0)}_{\beta})\,\psi^{(0)*}_{\beta} = \sum_{\beta} H^{(1)}_{\gamma\beta}\psi^{(0)*}_{\beta}.$$

By this, (5·11) can be written in the form

$$\sum_{\beta} H^{(1)}_{\gamma\beta}\mathscr{U}^{(1)}_{\beta\gamma} = N\int_{\Omega} d\boldsymbol{\rho}\,H^{(1)}\psi^{(0)*}_{\gamma}.\,D^{(1)}\psi^{(0)}_{\gamma} - NH^{(1)}_{\gamma\gamma}\int_{\Omega} d\boldsymbol{\rho}\,\psi^{(0)*}_{\gamma}.\,D^{(1)}\psi^{(0)}_{\gamma} + \sum_{\beta\neq\gamma} H^{(1)}_{\gamma\beta}X^{(1)}_{\beta\gamma}. \tag{5·12}$$

Substitute into this (5·7) for $H^{(1)}$ and transform the integral with the integrand $T^{(1)}\psi^{(0)*}_{\gamma}.\,D^{(1)}\psi^{(0)}_{\gamma}$ into another with the integrand $\psi^{(0)*}_{\gamma}.\,T^{(1)}D^{(1)}\psi^{(0)}_{\gamma}$ by partial integrations described at the beginning of this section. Then (5·12) becomes

$$\begin{aligned}\sum_{\beta} H^{(1)}_{\gamma\beta}\mathscr{U}^{(1)}_{\beta\gamma} = {} & N\int_{\Omega} d\boldsymbol{\rho}\,\psi^{(0)*}_{\gamma}.\,H^{(1)}D^{(1)}\psi^{(0)}_{\gamma} - NH^{(1)}_{\gamma\gamma}\int_{\Omega} d\boldsymbol{\rho}\,\psi^{(0)*}_{\gamma}.\,D^{(1)}\psi^{(0)}_{\gamma} + \sum_{\beta\neq\gamma} H^{(1)}_{\gamma\beta}X^{(\,)}_{\beta\gamma} \\ & + \frac{\hbar^2}{2m}\sum_{i,j=1}^{3} 2s_{ij}N\int_{\Omega} d\sigma\, n_i\left(\frac{\partial\psi^{(0)*}_{\gamma}}{\partial\rho_j}D^{(1)}\psi^{(0)}_{\gamma} - \psi^{(0)*}_{\gamma}\frac{\partial D^{(1)}\psi^{(0)}_{\gamma}}{\partial\rho_j}\right).\end{aligned} \tag{5·13}$$

When this is substituted into (5·9), we combine the integral of

$$\psi^{(0)*}_{\gamma}.\,(T^{(1)}+V^{(1)})\,D^{(1)}\psi^{(0)}_{\gamma}$$

with the integral of $\psi^{(0)*}_{\gamma}.\,(T^{(2)}+V^{(2)})\,\psi^{(0)}_{\gamma}$ and evaluate the combined integral by (3·11) and (5·1), putting in the latter $\beta=\gamma$ and $i=2$. Then we have for the second-order lattice energy

$$\begin{aligned}\Phi^{(2)} = \Phi^{(2)\prime} + \text{real part of } 2\sum_{\gamma=1}^{n}\Bigg\{ & N\int_{\Omega} d\boldsymbol{\rho}\,\psi^{(0)*}_{\gamma}\left(\sum_{l}^{\infty}{}' V^{(1)}_{l} + 2\sum_{\lambda=1}^{n}\mathscr{G}^{(1)}_{\lambda\lambda}\right)D^{(1)}\psi^{(0)}_{\gamma} \\ & - \left[E^{(1)}_{\gamma} + N\int_{\Omega} d\boldsymbol{\rho}\,\psi^{(0)*}_{\gamma}.\left(\sum_{l}^{\infty}{}' V^{(1)}_{l} + 2\sum_{\lambda=1}^{n}\mathscr{G}^{(1)}_{\lambda\lambda}\right)\psi^{(0)}_{\gamma}\right]N\int_{\Omega} d\boldsymbol{\rho}\,\psi^{(0)*}_{\gamma}.\,D^{(1)}\psi^{(0)}_{\gamma} \\ & + \sum_{\beta\neq\gamma} H^{(1)}_{\gamma\beta}X^{(1)}_{\beta\gamma} + \frac{\hbar^2}{2m}\sum_{i,j=1}^{3} 2s_{ij}N\int_{\Omega} d\sigma\, n_i\left(\frac{\partial\psi^{(0)*}_{\gamma}}{\partial\rho_j}D^{(1)}\psi^{(0)}_{\gamma} - \psi^{(0)*}_{\gamma}\frac{\partial D^{(1)}\psi^{(0)}_{\gamma}}{\partial\rho_j}\right) \\ & + \frac{\hbar^2}{2m}\sum_{i=1}^{3} N\int_{\Omega} d\sigma\, n_i\left(\psi^{(0)*}_{\gamma}\frac{\partial D^{(2)}\psi^{(0)}_{\gamma}}{\partial\rho_i} - \frac{\partial\psi^{(0)*}_{\gamma}}{\partial\rho_i}D^{(2)}\psi^{(0)}_{\gamma}\right) \\ & + N\int_{\Omega} d\boldsymbol{\rho}\,(V^{(0)\prime}_{\infty}\psi^{(0)*}_{\gamma}.\,D^{(2)}\psi^{(0)}_{\gamma} - \psi^{(0)*}_{\gamma}.\,D^{(2)}V^{(0)\prime}_{\infty}\psi^{(0)}_{\gamma})\Bigg\},\end{aligned} \tag{5·14}$$

where $\Phi^{(2)\prime}$ and $E^{(1)}_{\gamma}$ are given respectively by (5·10) and (5·6).

In (5·6) and (5·14) the interaction between an ion and an electron in the same atomic polyhedron is entirely eliminated. This avoids much numerical calculation necessitated by the fact that for short distances between the ion and the electron the interaction is known only numerically.

6. PERTURBATION ENERGY IN THE WIGNER AND SEITZ'S APPROXIMATION

The equations (5·5) and (5·14) give $\Phi^{(1)}$ and $\Phi^{(2)}$ as accurately as they can be obtained by the method of self-consistent field, if the unperturbed wave functions $\psi_\gamma^{(0)}$ are known to the same order of accuracy. At present, however, the $\psi_\gamma^{(0)}$ are only roughly known by the method of Wigner & Seitz (1933, 1934).

In order to get an approximate solution of the Fock-Dirac system of equations corresponding to the equilibrium configuration of the ions, these authors make use of the facts (i) that the metallic electrons of monovalent metals behave largely like free electrons, and (ii) that the atomic polyhedron for body-centred or face-centred cubic lattice approximates closely to a sphere.

Now it is known that the plane-wave functions are the eigenfunctions of the ordinary potential, as well as the exchange potential, that arises from an interaction dependent only on the relative co-ordinates of the interacting particles (Brillouin 1934; Seitz 1935). For an arbitrary interaction f of the relative co-ordinates $\boldsymbol{\rho}-\boldsymbol{\rho}'$, the eigenfunctions of the ordinary potential $P(f)$ and those of the exchange potential $A(f)$ are to be determined respectively from the equations

$$P(f)\,\psi(\boldsymbol{\rho}) = 2\sum_{\lambda=1}^{n}\int_{\infty} d\boldsymbol{\rho}'\,\psi_\lambda^*(\boldsymbol{\rho}')\,f(\boldsymbol{\rho}-\boldsymbol{\rho}')\,\psi_\lambda(\boldsymbol{\rho}')\,\psi(\boldsymbol{\rho}) = \text{const.}\,\psi(\boldsymbol{\rho}), \qquad (6\cdot1)$$

$$A(f)\,\psi(\boldsymbol{\rho}) = \sum_{\lambda=1}^{n}\int_{\infty} d\boldsymbol{\rho}'\,\psi_\lambda^*(\boldsymbol{\rho}')\,f(\boldsymbol{\rho}-\boldsymbol{\rho}')\,\psi(\boldsymbol{\rho}')\,\psi_\lambda(\boldsymbol{\rho}) = \text{const.}\,\psi(\boldsymbol{\rho}), \qquad (6\cdot2)$$

which are, however, satisfied by plane-wave functions.

If we apply the above to the Coulomb interaction of the electrons, taking $f = \mathscr{G}$, we arrive at the result that, to the approximation of free electrons, the exchange and the ordinary potential of the electrons in monovalent metals are constant.

To the approximation that the basic atomic polyhedron is replaced by a sphere, the lattice arrangement of the ions outside is replaced by a spherical distribution. The potential energy of an electron inside the atomic sphere due to its interactions with the ions outside arises chiefly from the Coulomb interaction and is, by electrostatics, constant.

In our notation the above means that $V_\infty^{(0)\prime}$ is constant. From (4·14) and (4·8) we then arrive at the wave equations of Wigner & Seitz:

$$(T+V-E_\gamma^{(0)})\,\psi_\gamma^{(0)}=0, \quad \text{or} \quad \left\{-\frac{\hbar^2}{2m}(\nabla^2+2i\mathbf{k}\,.\,\nabla-k^2)+V-E_\gamma^{(0)}\right\}u_\gamma^{(0)}=0, \tag{6·3}$$

where the parameters $E_\gamma^{(0)}$ of Wigner & Seitz differ from the $W_\gamma^{(0)}$ of Fock & Dirac merely by the constant $V_\infty^{(0)\prime}$. The $u_\gamma^{(0)}$ are normalized in the atomic sphere if we choose this to be of the same volume as the atomic polyhedron. We shall refer to $E_\gamma^{(0)}$ as the Wigner-Seitz energy.

By making use of (6·3), the potential energy of the lattice in equilibrium is given by

$$\Phi^{(0)}=2\sum_{\gamma=1}^{n}E_\gamma^{(0)}+2\sum_{\gamma=1}^{n}\sum_{l}^{\infty}{}'N\int_\Omega d\boldsymbol{\rho}\,\psi_\gamma^{(0)*}V_l^{(0)}\psi_\gamma^{(0)}$$
$$+2\sum_{\gamma,\lambda=1}^{n}\mathscr{G}_{\gamma\lambda\lambda\gamma}^{(0)}-\sum_{\gamma,\lambda=1}^{n}\mathscr{G}_{\gamma\lambda\gamma\lambda}^{(0)}+\tfrac{1}{2}\sum_{l}\sum_{l'}^{\infty}\phi_{ll'}^{(0)}. \tag{6·4}$$

We now consider the perturbation energy due to a homogeneous deformation, assuming that the unperturbed system of the Fock-Dirac equations can be solved by the approximate method of Wigner & Seitz. We have then in $\boldsymbol{\rho}$-space a spherical distribution of 'ions' outside the atomic sphere, the interaction of an 'ion' and an 'electron' in $\boldsymbol{\rho}$-space being, however, $\overline{V}_l$. In order to show that the potential energy of an 'electron' inside the atomic sphere arising from its interactions $\overline{V}_l$ with the spherical distribution of 'ions' outside is also constant, we apply the inverse transformation of (4·1), i.e. from $\boldsymbol{\rho}$ to $\mathbf{r}$. This transforms the family of concentric spheres in $\boldsymbol{\rho}$-space into the family of concentric and similar ellipsoids in $\mathbf{r}$-space, and *the interaction* $\overline{V}_l=V(\mathbf{r}(\boldsymbol{\rho}),\mathbf{r}(\mathbf{a}^l))$ *into the Coulomb interaction* $V(\mathbf{r},\mathbf{r}^l)$. We can then apply the theorem of the potential theory that the potential inside the cavity of a uniformly charged shell bounded by concentric and similar ellipsoidal surfaces is constant (e.g. Moulton 1902, p. 92).

By applying (6·1) and (6·2) to the interaction $\overline{\mathscr{G}}$ of the 'electrons' in $\boldsymbol{\rho}$-space, taking the unperturbed plane-wave functions for the ψ's, we see that the ordinary and the exchange potentials arising from $\overline{\mathscr{G}}(\boldsymbol{\rho}-\boldsymbol{\rho}')$ are also constant.

The above consideration holds for an arbitrary homogeneous deformation. Hence we have

$$\sum_{l}^{\infty}{}'V_l^{(i)}+2\sum_{\lambda=1}^{n}\mathscr{G}_{\lambda\lambda}^{(i)}=\text{constant}, \quad \sum_{\lambda=1}^{n}\mathscr{G}_{\beta\lambda\gamma\lambda}^{(i)}=0 \quad (\beta\neq\gamma) \quad (i=0,1,\ldots). \tag{6·5}$$

With $i=0$, (6·5) implies again that $V_\infty^{(0)\prime}$ is constant. We use this and (6·5) with $i=1$ to simplify the expressions (5·6) and (5·14) for the perturbation energy.

Then we get

$$\Phi^{(i)} = 2\sum_{\gamma=1}^{n} E_{\gamma}^{(i)} + 2\sum_{\gamma=1}^{n}\Big|\sum_{l}^{\infty}{}' N\int_{\Omega} d\boldsymbol{\rho}\, \psi_{\gamma}^{(0)*} V_{l}^{(i)} \psi_{\gamma}^{(0)} + 2\sum_{\gamma,\,\lambda=1}^{n} \mathscr{G}_{\gamma\lambda\lambda\gamma}^{(i)} - \sum_{\gamma,\,\lambda=1}^{n} \mathscr{G}_{\gamma\lambda\gamma\lambda}^{(i)} + \tfrac{1}{2}\sum_{l}\sum_{l'}^{\infty} \phi_{ll'}^{(i)}, \tag{6·6}$$

with
$$E_{\gamma}^{(1)} = \frac{\hbar^2}{2m}\sum_{i=1}^{3} N\int_{\Omega} d\sigma\, n_i\left(\psi_{\gamma}^{(0)*}\frac{\partial D^{(1)}\psi_{\gamma}^{(0)}}{\partial\rho_i} - \frac{\partial\psi_{\gamma}^{(0)*}}{\partial\rho_i} D^{(1)}\psi_{\gamma}^{(0)}\right), \tag{6·7}$$

and (we use for the $X_{\beta\gamma}^{(1)}$ of (5·4) the zero approximation given by (5·3))

$$\begin{aligned} E_{\gamma}^{(2)} = \text{real part of } \Bigg\{ & \frac{\hbar^2}{2m}\sum_{i=1}^{3} N\int_{\Omega} d\sigma\, n_i\left(\psi_{\gamma}^{(0)*}\frac{\partial D^{(2)}\psi_{\gamma}^{(0)}}{\partial\rho_i} - \frac{\partial\psi_{\gamma}^{(0)*}}{\partial\rho_i} D^{(2)}\psi_{\gamma}^{(0)}\right) \\ & + \frac{\hbar^2}{2m}\sum_{i,\,j=1}^{3} 2S_{ij} N\int_{\Omega} d\sigma\, n_i\left(\frac{\partial\psi_{\gamma}^{(0)*}}{\partial\rho_j} D^{(1)}\psi_{\gamma}^{(0)} - \psi_{\gamma}^{(0)*}\frac{\partial D^{(1)}\psi_{\gamma}^{(0)}}{\partial\rho_j}\right) \\ & - E_{\gamma}^{(1)} N\int_{\Omega} d\boldsymbol{\rho}\, \psi_{\gamma}^{(0)*} D^{(1)}\psi_{\gamma}^{(0)} + \sum_{\beta\neq\gamma} N\int_{\Omega} d\boldsymbol{\rho}\, \psi_{\gamma}^{(0)*}\left(T^{(1)} + V^{(1)}\right)\psi_{\beta}^{(0)} \frac{1}{W_{\beta}^{(0)} - W_{\gamma}^{(0)}} \\ & \times \frac{\hbar^2}{2m}\sum_{i=1}^{3} N\int_{\Omega} d\sigma\, n_i\left(\frac{\partial\psi_{\beta}^{(0)*}}{\partial\rho_i} D^{(1)}\psi_{\gamma}^{(0)} - \psi_{\beta}^{(0)*}\frac{\partial D^{(1)}\psi_{\gamma}^{(0)}}{\partial\rho_i}\right). \end{aligned} \tag{6·8}$$

On comparing (6·6) with (6·4) we notice the similar structure of these expressions for $\Phi^{(0)}$, $\Phi^{(1)}$, $\Phi^{(2)}$. The fact that $E_{\gamma}^{(1)}$ and $E_{\gamma}^{(2)}$ are independent of the mutual interaction between the electrons shows clearly that they represent the perturbation of the Wigner-Seitz energy. The perturbation to the wave equations of Wigner & Seitz is merely a change of the boundary arising from the deformation of an atomic sphere into an ellipsoid. Thus $E_{\gamma}^{(1)}$ and $E_{\gamma}^{(2)}$ are composed of surface integrals.

We shall now evaluate the above expressions for $E_{\gamma}^{(1)}$ and $E_{\gamma}^{(2)}$.

For the alkali metals—with the exception of lithium—and also for copper and silver, it is known that the wave functions for the occupied states are given in fair approximation by

$$\psi_{\gamma}^{(0)} = N^{-\frac{1}{2}} e^{i\mathbf{k}\cdot\boldsymbol{\rho}} u(\rho), \tag{6·9}$$

where $u(\rho)$ does not depend on $\mathbf{k}$. With such wave functions the Wigner-Seitz energy $E_{\gamma}^{(0)}$ is given by

$$E_{\gamma}^{(0)} = E^{(0)} + \frac{\hbar^2}{2m} k^2, \tag{6·10}$$

where $E^{(0)}$ is the Wigner-Seitz energy of the lowest state (with $\mathbf{k} = 0$). Being a function of the distance only, $u(\rho)$ satisfies the equation

$$\left\{-\frac{\hbar^2}{2m}\left(\frac{d^2}{d\rho^2} + \frac{2}{\rho}\frac{d}{d\rho}\right) + V(\rho) - E^{(0)}\right\} u(\rho) = 0 \tag{6·11}$$

with the boundary condition

$$(du/d\rho)_{\rho=r_0} = 0, \tag{6·12}$$

r_0 being the radius of the atomic sphere, i.e. $4\pi r_0^3/3 = \Omega$, the atomic volume of the undeformed lattice. From (6·11) and (6·12) follow

$$\frac{\hbar^2}{2m}\left(\frac{d^2u}{d\rho^2}\right)_{\rho=r_0} = \{V(r_0) - E^{(0)}\}\,u(r_0) \tag{6·13}$$

and

$$\frac{\hbar^2}{2m}\left(\frac{d^3u}{d\rho^3}\right)_{\rho=r_0} = \left[\left(\frac{dV}{d\rho}\right)_{\rho=r_0} - \frac{2}{r_0}\{V(r_0) - E^{(0)}\}\right]u(r_0). \tag{6·14}$$

Further, we denote by f the quantity

$$f = \frac{4\pi r_0^3}{3}\,u^2(r_0). \tag{6·15}$$

Substituting (6·9) for $\psi_\gamma^{(0)}$ and (4·4) for $D^{(1)}$ and $D^{(2)}$ into (6·7) and (6·8) and making use of (6·12)–(6·15), we obtain

$$E_\gamma^{(1)} = \sum_{i,j=1}^{3} f s_{ij}\left[\delta_{ij}\{V(r_0) - E^{(0)}\} - \frac{\hbar^2}{2m}\,2k_i k_j\right] \tag{6·16}$$

and

$$\begin{aligned} E_\gamma^{(2)} = \sum_{i,j,p,q=1}^{3} f s_{ij} s_{pq} \Big\{ & \tfrac{3}{2} n_{ijpq}\left[r_0\left(\frac{dV}{d\rho}\right)_{\rho=r_0} - 5\{V(r_0) - E^{(0)}\}\right] + \tfrac{1}{2}\delta_{iq}\delta_{jp} \\ & \times \{V(r_0) - E^{(0)}\} + \frac{\hbar^2}{2m} 3k_i \delta_{jp} k_q - \frac{f-1}{2}\delta_{ij}\left[\delta_{pq}\{V(r_0) - E^{(0)}\} - \frac{\hbar^2}{2m} 2k_p k_q\right]\Big\} \\ & + \text{real part of} \sum_{\beta \neq \gamma} N \int_\Omega d\rho\, \psi_\gamma^{(0)*} \,.\, (T^{(1)} + V^{(1)})\, \psi_\beta^{(0)} \frac{1}{W_\beta^{(0)} - W_\gamma^{(0)}} \\ & \times \frac{\hbar^2}{2m} \sum_{i=1}^{3} N \int_\Omega d\sigma\, n_i \left(\frac{\partial \psi_\beta^{(0)*}}{\partial \rho_i} D^{(1)} \psi_\gamma^{(0)} - \psi_\beta^{(0)*} \frac{\partial D^{(1)} \psi_\gamma^{(0)}}{\partial \rho_i}\right). \end{aligned} \tag{6·17}$$

Here n_{ijpq} denotes the average of the product $n_i n_j n_p n_q$, viz. $n_{1111} =$ etc. $= 1/5$, $n_{1122} = n_{1212} =$ etc. $= 1/15$, and $n_{1123} =$ etc. $= 0$.

The last line of (6·17) cannot be calculated without the knowledge of the eigenfunctions $\psi_\beta^{(0)}$ and eigenvalues $W_\beta^{(0)}$ of the excited bands. For simplicity, we shall neglect it. This implies that the factor f is treated as constant, as will be shown below.

We consider the perturbation of the Wigner-Seitz energy of the lowest state $\mathbf{k} = 0$ due to a uniform dilatation

$$s_{ij} = \delta_{ij}(\Delta r_0/r_0). \tag{6·18}$$

The perturbation expansion for the Wigner-Seitz energy of the lowest state (we shall omit the index γ for $\mathbf{k} = 0$)

$$E = E^{(0)} + E^{(1)} + E^{(2)} + \dots \tag{6·19}$$

then must coincide with the Taylor expansion

$$E(r_0+\Delta r_0) = E(r_0)+\Delta r_0\frac{dE}{dr_0}+\tfrac{1}{2}(\Delta r_0)^2\frac{d^2E}{dr_0^2}+\dots, \tag{6·20}$$

where E is considered as a function of the radius of the atomic sphere. Substituting (6·18) into (6·16) and (6·17) we obtain, by comparison of (6·19) and (6·20), the relations $E(r_0) = E^{(0)}$,

$$dE/dr_0 = 3f\{V(r_0)-E^{(0)}\}/r_0, \tag{6·21}$$

and

$$\begin{aligned}\frac{1}{2}\frac{d^2E}{dr_0^2} = \frac{3f}{2}&\left\{\frac{1}{r_0}\left(\frac{dV}{d\rho}\right)_{\rho=r_0}-4\frac{V(r_0)-E^{(0)}}{r_0^2}-3(f-1)\frac{V(r_0)-E^{(0)}}{r_0^2}\right\}\\&+\frac{V(r_0)-E^{(0)}}{r_0}\sum_\beta{}'\int_\Omega d\rho\, u\left(\frac{\hbar^2}{m}\nabla^2+\rho\frac{dV}{d\rho}\right)u_\beta^{(0)}\frac{1}{E^{(0)}-E_\beta^{(0)}}\int_\Omega d\sigma\, u_\beta^{(0)}u,\end{aligned} \tag{6·22}$$

the summation covering the unoccupied states with $\mathbf{k} = 0$.

The relation (6·21) has been derived by Bardeen (1938) by Fröhlich's method of 'perturbation of boundary conditions' (1937, 1938), while (6·22) is evidently its analogue. Forming d^2E/dr_0^2 from (6·21) and comparing this with (6·22), we get, after removing the common factor $\{V(r_0)-E^{(0)}\}/r_0$ from both sides,

$$\frac{3}{2}\frac{df}{dr_0} = \sum_\beta{}'\int_\Omega d\rho\, u\left(\frac{\hbar^2}{m}\nabla^2+\rho\frac{dV}{d\rho}\right)u_\beta^{(0)}\frac{1}{E^{(0)}-E_\beta^{(0)}}\int_\Omega d\sigma\, u_\beta^{(0)}u. \tag{6·23}$$

This shows that if we neglect the last line of (6·17) we can then treat f as constant. Thanks to previous numerical calculations (Wigner & Seitz (1933, 1934) and Bardeen (1938) for sodium; Fuchs (1935) for copper and silver), we shall take in the following $f = 1$. On the right-hand side of (6·17) the term containing the factor $(f-1)$ then vanishes. We rewrite, therefore, (6·16) and (6·17) as follows:

$$E_\gamma^{(1)} = \sum_{i,j=1}^{3}\delta_{ij}s_{ij}\{V(r_0)-E^{(0)}\}-\frac{\hbar^2}{2m}\sum_{i,j=1}^{3}2s_{ij}k_ik_j, \tag{6·24}$$

$$\begin{aligned}E_\gamma^{(2)} = \sum_{i,j,p,q=1}^{3}s_{ij}s_{pq}&\left\{\tfrac{3}{2}n_{ijpq}\left[r_0\left(\frac{dV}{d\rho}\right)_{\rho=r_0}-5\{V(r_0)-E^{(0)}\}\right]\right.\\&\left.+\tfrac{1}{2}\delta_{iq}\delta_{jp}\{V(r_0)-E^{(0)}\}\right\}+\frac{\hbar^2}{2m}\sum_{i,j,q}3s_{ij}s_{jq}k_ik_q.\end{aligned} \tag{6·25}$$

The dependence of the Wigner-Seitz energy $E_\gamma^{(i)}$ $(i = 0,1,2,\dots)$ on the vector $\mathbf{k}$ as given by (6·10), (6·24) and (6·25) can be summarized in the single equation

$$E_\gamma = E+\frac{\hbar^2}{2m}\bar{\mathbf{k}}^2, \tag{6·26}$$

where $E = E^{(0)}+E^{(1)}+E^{(2)}+\dots$, $E_\gamma = E_\gamma^{(0)}+E_\gamma^{(1)}+E_\gamma^{(2)}+\dots$ and $\bar{\mathbf{k}}$ is simply the wave vector in the deformed lattice. Neglecting any temperature effects, it follows that all states within a sphere of radius $\bar{k}_0$ in $\bar{\mathbf{k}}$-space are occupied, and the total Wigner-Seitz energy of all the electrons is given by

$$2\sum_{\gamma=1}^{n} E_\gamma = NE+NF. \tag{6·27}$$

Here F is the average 'Fermi energy' (Mott & Jones 1936, pp. 54, 81)

$$F = \frac{3}{5}\frac{\hbar^2}{2m}\left(\frac{3\pi^2}{\bar{\Omega}}\right)^{\frac{2}{3}} = \frac{3}{10}\left(\frac{9\pi}{4}\right)^{\frac{2}{3}}\frac{\hbar^2}{m\bar{r}_0^2}, \tag{6·28}$$

where $\bar{\Omega}$, $= 4\pi\bar{r}_0^3/3$, denotes the atomic volume of the deformed lattice.

7. THE ELASTIC CONSTANTS

The total lattice energy is

$$\Phi = \Phi^{(0)}+\Phi^{(1)}+\Phi^{(2)}+\dots, \tag{7·1}$$

where $\Phi^{(0)}$, $\Phi^{(1)}$, $\Phi^{(2)}$ are given by (6·4) and (6·6). With the help of (6·27) and (2·4) we can write (7·1) in the form

$$\Phi = NE+NF+2\sum_{\gamma=1}^{n}\sum_{l}^{\infty}{}'\int_\Omega d\boldsymbol{\rho}\,\psi_\gamma^{(0)*}\bar{V}_l\psi_\gamma^{(0)} \\ +2\sum_{\gamma,\lambda=1}^{n}\bar{\mathscr{G}}_{\gamma\lambda\lambda\gamma}-\sum_{\gamma,\lambda=1}^{n}\bar{\mathscr{G}}_{\gamma\lambda\gamma\lambda}+\tfrac{1}{2}\sum_{l}\sum_{l'}^{\infty}\bar{\Phi}_{ll'}. \tag{7·2}$$

We shall split (7·2) into the terms

$$= N(E+F+A+L-S+I). \tag{7·3}$$

Here E is the energy of the lowest state and F the Fermi energy; together they give the Wigner-Seitz energy, which has been considered in the last section.

NA is given by
$$NA = -\sum_{\gamma,\lambda=1}^{n}\bar{\mathscr{G}}_{\gamma\lambda\gamma\lambda} \tag{7·4}$$

and represents the exchange energy between electrons with parallel spins.

NL is the lattice energy of an electrically neutral system composed of an array of positive point charges plus a uniform, continuous distribution of negative charges, in $\mathbf{r}$-space:

$$NL = \frac{1}{2}\int\frac{d\boldsymbol{\rho}}{\Omega}\int_\infty\frac{d\boldsymbol{\rho}'}{\Omega}\frac{e^2}{|\mathbf{r}(\boldsymbol{\rho})-\mathbf{r}(\boldsymbol{\rho}')|}-\sum_{l}^{\infty}\int\frac{d\boldsymbol{\rho}}{\Omega}\frac{e^2}{|\mathbf{r}(\boldsymbol{\rho})-\mathbf{r}^l|}+\tfrac{1}{2}\sum_{l}\sum_{l'}^{\infty}\frac{e^2}{|\mathbf{r}^l-\mathbf{r}^{l'}|}. \tag{7·5}$$

As discussed by other authors (Wigner & Seitz 1934; Fuchs 1936), the actual interaction between electrons and ions in a metal can be reasonably represented by such an idealized model, with the exception of the interaction between electron and ion in the same atomic polyhedron. We can also write (7·5) in the form

$$NL = 2\sum_{\gamma,\lambda=1}^{n} \overline{\mathscr{G}}_{\gamma\lambda\lambda\gamma} + 2\sum_{\gamma=1}^{n}\sum_{l}^{\infty}{}' N\int_{\Omega} \psi_{\gamma}^{(0)*}\,\overline{V}_l\,\psi_{\gamma}^{(0)} - N\int_{\Omega}\frac{d\boldsymbol{\rho}}{\Omega}\frac{e^2}{|\mathbf{r}(\boldsymbol{\rho})|} + \tfrac{1}{2}\sum_{l}\sum_{l'}^{\infty}\frac{e^2}{|\mathbf{r}^l - \mathbf{r}^{l'}|}, \tag{7·6}$$

remembering the approximations used (viz. $\overline{V}_l = -e^2/|\mathbf{r}(\boldsymbol{\rho}) - \mathbf{r}^l|$ when l_1, l_2, l_3 do not all vanish and $\psi_{\gamma}^{(0)} =$ plane waves $= (N\Omega)^{-\frac{1}{2}}\,.\,e^{i\mathbf{k}.\boldsymbol{\rho}}$ for $\gamma = 1, 2, \ldots, n$).

The first term of (7·6) gives the Coulomb interaction between the electrons; the second term, the interaction between electrons and ions with the exception of electron and ion in the same atomic polyhedron which is given, for the idealized model, by the third term; and the last term, the Coulomb interaction between the ions.

The third term of (7·6) has no significance for the actual metal; it has only been added, because then NL is a quantity that can conveniently be calculated. In order to obtain the actual lattice energy it must be subtracted again. This is done by means of the quantity NS defined by

$$NS = -N\int_{\Omega}\frac{d\boldsymbol{\rho}}{\Omega}\frac{e^2}{|\mathbf{r}(\boldsymbol{\rho})|}. \tag{7·7}$$

Further, the interaction between the ions contains not only the Coulomb interaction, which is taken care of in NL, but also the van der Waals forces and repulsive forces. These are collected in the term NI of (7·3),

$$NI = \tfrac{1}{2}\sum_{l}\sum_{l'}^{\infty}\{\phi(\mathbf{r}^l, \mathbf{r}^{l'}) - e^2/|\mathbf{r}^l - \mathbf{r}^{l'}|\}. \tag{7·8}$$

Comparison of (7·4), (7·6), (7·7) and (7·8) with (7·2) shows that we have taken into account all terms.

We consider first the exchange energy (7·4). By applying the co-ordinate transformation from $\boldsymbol{\rho}$ to $\mathbf{r}$, we have

$$NA = -\int\frac{d\mathbf{r}}{\overline{\Omega}}\int_{\infty}\frac{d\mathbf{r}'}{\overline{\Omega}}\frac{e^2}{|\mathbf{r}-\mathbf{r}'|}\sum_{\overline{\mathbf{k}},\overline{\mathbf{k}}'} e^{i(\overline{\mathbf{k}}-\overline{\mathbf{k}}').(\mathbf{r}-\mathbf{r}')}, \tag{7·9}$$

where the summation over $\overline{\mathbf{k}}$ and $\overline{\mathbf{k}}'$ extends over all the occupied states, i.e. all the states within a sphere of radius $\overline{k}_0$ in $\overline{\mathbf{k}}$-space. This has been evaluated by Wigner & Seitz with the result

$$A = -0{\cdot}458e^2/\overline{r}_0. \tag{7·10}$$

This depression of energy shows that electrons in metals possess a tendency to keep away from each other when their spins are parallel. Because of the plane-wave functions used for the $\psi_\gamma^{(0)}$, A depends only on the atomic volume.

In their investigation on the binding of sodium metal Wigner & Seitz (1934) have shown that there is also a depression of the energy owing to the tendency of the electrons with antiparallel spins to keep away from each other. This tendency is not taken into account in the method of self-consistent field. Wigner (1934) therefore introduces an extension of this method; the extension enables one to calculate the extra energy due to the correlation of the positions of electrons with antiparallel spins after one has solved the Fock-Dirac equations. This extra energy has been called 'correlation energy'. In case the Fock-Dirac equations are solved by plane-wave functions, the correlation energy NC is a function of the atomic volume only. C is given by (Wigner 1934, figure 7, p. 1010)

$$C = -\frac{e^2}{\bar{r}_0}\left(0{\cdot}0170 + 0{\cdot}0295\,\frac{\bar{r}_0}{a_0} - 0{\cdot}0015\,\frac{\bar{r}_0^2}{a_0^2}\right), \quad 1 < \frac{\bar{r}_0}{a_0} < 8, \qquad (7{\cdot}11)$$

where a_0 denotes the atomic unit of length, i.e. $a_0 = \hbar^2/me^2$. Hence, to the approximation of Wigner, we have to add the term NC to the right-hand side of (7·3). Its contribution to the elastic constants will be seen in the following paragraphs:

(*a*) *Elastic deformation without change of volume.* We consider now an elastic deformation without change of volume. It has been shown above that F, A and C depend only on the atomic volume ((6·28), (7·10) and (7·11)). In the appendix it will be shown that in addition $E-S$ depends only on the volume. Thus the two independent combinations of elastic constants corresponding to elastic deformations without change of volume are given by the variation of $N(L+I)$, i.e. by the variation of the Coulomb energy of the idealized model and the non-Coulomb energy of the ions.

The two relevant combinations of the elastic constants have been calculated previously (Fuchs 1936) from the variation of $N(L+I)$. At the time the assumption that this is sufficient could only be made a probable one. We can now see exactly to what extent the approximation is justified. Indeed, the approximation is rather better than could be expected at the time. For most monovalent metals the approximation is quite sufficient. The only exception is lithium, for which better unperturbed wave functions than (6·9) are to be used.

(*b*) *The bulk modulus.* The bulk modulus has been calculated previously by one of us from the variation of $E+F+I$ and by other authors quoted at the beginning of the Introduction. In the approximations $f=1$, $df/dr_0=0$, we can eliminate the quantity $\{V(r_0)-E^{(0)}\}$, which arises from both dE/dr_0 and d^2E/dr_0^2 (cf. (6·21), (6·22)), between the equation of state and the expression for the bulk modulus. We obtain thus for the bulk modulus $1/\kappa$ (κ being the compressibility) at zero pressure

$$\frac{1}{\kappa}=\frac{1}{9\Omega}r_0^2\frac{d^2\Phi/N}{dr_0^2}=\frac{1}{4\pi r_0^3}\left\{\frac{e^2}{r_0}-\tfrac{2}{3}F+\left(\tfrac{4}{3}r_0\frac{d}{dr_0}+\tfrac{1}{3}r_0^2\frac{d^2}{dr_0^2}\right)(I+A+L-S+C)\right\}, \tag{7·12}$$

which differ from the previous formula (Fuchs 1936*a*, p. 445) merely by the additional contribution arising from the variation of the term $A+L-S+C$ of the lattice energy.

Substituting for A and C the values (7·10) and (7·11) (of course with r_0 now in place of $\bar{r}_0$), for L, the values $L=-0{\cdot}89586e^2/r_0$ (for face-centred cubic lattice) and $L=-0{\cdot}89593e^2/r_0$ (for body-centred cubic lattice) calculated previously (Wigner & Seitz 1934; Fuchs 1935) and for S the value $S=-\,3e^2/2r_0$ given at the end of the Appendix, we find

$$\frac{1}{\kappa}=\left(\frac{1}{\kappa}\right)_{EFI}-0{\cdot}097\frac{e^2}{4\pi r_0^4}+\left(0{\cdot}011+0{\cdot}002\frac{r_0^2}{a_0^2}\right)\frac{e^2}{4\pi r_0^4}. \tag{7·13}$$

Here, on the right, the first term $(1/\kappa)_{EFI}$ arises from the variation of $E+F+I$, the second term, from that of $A+L-S$, and the third term, from C, the difference of L for the face-centred and body-centred lattices being non-significant.

In the following table we compare, on the one hand, the values for the bulk modulus $1/\kappa$ (in 10^{11} c.g.s. units) calculated to various approximations with the experimental values and, on the other hand, the various calculated values for the bulk modulus of sodium with those for the binding energy $\Phi^{(0)}$ (in kcal./mol.) of sodium:

	terms of Φ included to the approximation			
	$E+F+I$	$E+F+I+A$ $+L-S$	$E+F+I+A$ $+L-S+C$	experimental
$1/\kappa$ of Li*	1·76	1·54	1·61	1·30†
$1/\kappa$ of Na	0·878†	0·792	0·820	0·813§
($\Phi^{(0)}$ of Na)‡	(25·6)	(9·0)	(26·1)	(26·9)
$1/\kappa$ of K	0·406†	0·363	0·388	0·40†
$1/\kappa$ of Cu	14·1†	13·7	13·8	13·9†

* For lithium the approximation (6·9) is not justified.

† These values are from Fuchs (1936*a*) (for alkalis) and (1935) (for Cu). The experimental values refer to zero temperature, extrapolated by Fuchs.

‡ The values in this row are from Wigner & Seitz (1933) and Wigner (1934).

§ This refers to zero temperature and zero pressure, obtained by Bardeen (1938) by extrapolation. His calculated value for $1/\kappa$ of Na is 0·833 in the above units. In his calculation, he neglected the variation of I, the non-Coulomb energy of the ions, but, instead of taking $f=0$ and $df/dr_0=0$, he calculated the values of f for a range of values of r_0 by numerical methods.

We see that the second and third terms of (7·13) definitely improve the agreement between the calculated and experimental values of the bulk modulus, but the correlation energy C plays a much less important role in the calculation of the elastic constants than it does in the calculation of the binding energy of metals.

We wish to thank Professor Max Born for the interest he has taken in the present work.

APPENDIX

We shall now prove that the combination $E-S$ contributes nothing to the elastic resistance at constant volume.

For cubic lattices the perturbation expansion

$$\Phi = \Phi^{(0)}+\Phi^{(1)}+\Phi^{(2)}+\ldots = \Phi^{(0)}+\sum_{i,j=1}^{3} s_{ij}\Phi^{ij}+\sum_{i,j,p,q=1}^{3} s_{ij}s_{pq}\Phi^{ijpq}+\ldots$$

assumes necessarily the form

$$\Phi = \Phi^{(0)}+\Phi_1(s_{11}+s_{22}+s_{33})+\frac{\Phi_{11}}{2}(s_{11}^2+s_{22}^2+s_{33}^2)$$
$$+\Phi_{12}(s_{11}s_{22}+s_{22}s_{33}+s_{33}s_{11})+2\Phi_{44}(s_{12}^2+s_{23}^2+s_{31}^2)+\ldots,$$

where Φ_1, Φ_{11}, Φ_{12}, Φ_{44} are given in terms of Φ^{ij}, Φ^{ijpq} by

$$\Phi_1 = \Phi^{11} = \text{etc.},\quad \Phi_{12} = \Phi^{1122}+\Phi^{2211} = \text{etc.},$$
$$\frac{\Phi_{11}}{2} = \Phi^{1111} = \text{etc.},\quad 2\Phi_{44} = \Phi^{1212}+\Phi^{2121}+\Phi^{1221}+\Phi^{2112} = \text{etc.} \tag{i}$$

When the volume is kept constant, there are only two independent elastic distortions which we choose as follows (Fuchs 1936):

(A) Compression and expansion parallel to two of the cube edges, the compression and the expansion being so related that the volume is unchanged.

(B) Uniform shear parallel to one surface of the cube.

Let ϵ_A and ϵ_B denote respectively the parameters measuring these distortions, ϵ_A being the fractional elongation of the cube edge and ϵ_B the angle of shear. The elastic resistances for these distortions are measured respectively by $\Phi_A = (1/2)\,d^2\Phi/d\epsilon_A^2$ and $\Phi_B = (1/2)\,d^2\Phi/d\epsilon_B^2$. These are given in terms of Φ_1, Φ_{11}, etc., by

$$\Phi_A = \Phi_1+\Phi_{11}-\Phi_{12},\quad \Phi_B = \frac{\Phi_1}{4}+\frac{\Phi_{44}}{2}. \tag{ii}$$

These relations, being linear in Φ, hold for each term of Φ, e.g. for E or S.

In order to prove that $E-S$ contributes nothing to the elastic resistance at constant volume, we have to show that $E_A = S_A$ and $E_B = S_B$. From (6·24) and (6·25), we obtain in accordance with (i)

$$E_1 = \{V(r_0)-E^{(0)}\},\quad E_{11} = \tfrac{3}{5}r_0\left(\frac{dV}{d\rho}\right)_{\rho=r_0}-2\{V(r_0)-E^{(0)}\},$$
$$E_{12} = \tfrac{1}{5}r_0\left(\frac{dV}{d\rho}\right)_{\rho=r_0}-\{V(r_0)-E^{(0)}\},\quad E_{44} = \tfrac{1}{5}r_0\left(\frac{dV}{d\rho}\right)_{\rho=r_0}-\tfrac{1}{2}\{V(r_0)-E^{(0)}\}.$$

Hence, by (ii), we have for E_A and E_B (V tends to $-e^2/\rho$ for large ρ)

$$E_A = \tfrac{2}{5}r_0\left(\frac{dV}{d\rho}\right)_{\rho=r_0} = \frac{2}{5}\frac{e^2}{r_0}, \quad E_B = \tfrac{1}{10}r_0\left(\frac{dV}{d\rho}\right)_{\rho=r_0} = \frac{1}{10}\frac{e^2}{r_0}. \qquad \text{(iii)}$$

They satisfy of course the isotropic condition $E_A = 4E_B$ as the atomic polyhedron has been approximated by a sphere.

For the calculation of S_A and S_B we need the quantities S^{ij} and S^{ijpq} which, by (7·7), are given by

$$S^{ij} = -\int_\Omega \frac{d\rho}{\Omega}\rho_i\frac{\partial}{\partial\rho_j}\left(\frac{e^2}{\rho}\right), \quad S^{ijpq} = -\int_\Omega \frac{d\rho}{\Omega}\rho_i\rho_p\frac{1}{2}\frac{\partial^2}{\partial\rho_j\partial\rho_q}\left(\frac{e^2}{\rho}\right). \qquad \text{(iv)}$$

Applying Euler's theorem on homogeneous functions we have from (iv)

$$\sum_{i=1}^{3} S^{ii} = -S^{(0)}, \qquad \text{(v)}$$

and

$$\sum_{i,p=1}^{3} S^{iipp} = \sum_{i,p=1}^{3} S^{ippi} = S^{(0)}, \qquad \text{(vi)}$$

while applying Poisson's equation we have from (iv)

$$\sum_{i,p=1}^{3} S^{ipip} = 0.$$

This last relation may be replaced in virtue of (vi) by

$$\sum_{i,p=1}^{3} S^{ipip} + \sum_{i,p=1}^{3} S^{ippi} = S^{(0)}. \qquad \text{(vii)}$$

The relations (v), (vi) and (vii), when written in terms of S_1, S_{11}, etc., become (cf. (i))

$$3S_1 = -S^{(0)}, \quad 3\left(\frac{S_{11}}{2}+S_{12}\right) = S^{(0)}, \quad 6\left(\frac{S_{11}}{2}+S_{44}\right) = S^{(0)}. \qquad \text{(viii)}$$

We have further the isotropic condition

$$S_A = 4S_B, \quad \text{i.e.} \quad S_{11}-S_{12} = 2S_{44}. \qquad \text{(ix)}$$

These four equations can be solved for S_1, S_{11}, S_{12}, S_{44} in terms of $S^{(0)}$ and yield thus for S_A and S_B

$$S_A = -\tfrac{4}{15}S^{(0)}, \quad S_B = -\tfrac{1}{15}S^{(0)}. \qquad \text{(x)}$$

By direct integration we have

$$S^{(0)} = -\int_\Omega \frac{d\rho}{\Omega}\frac{e^2}{\rho} = -\frac{3e^2}{2r_0}.$$

Substituting this value into (x) and comparing the result with (iii), we see that $E_A = S_A$ and $E_B = S_B$. Q.E.D.

REFERENCES

Bardeen 1938 *J. Chem. Phys.* **6**, 367, 372.
Bloch 1928 *Z. Phys.* **52**, 585.
Born 1923 *Atomtheorie des festen Zustandes.* Leipzig.
Brillouin 1934 *J. Phys. Radium*, **5**, 413.
Fröhlich 1937 *Proc. Roy. Soc.* A, **158**, 97.
Fröhlich 1938 *Phys. Rev.* **54**, 945.
Fuchs 1935 *Proc. Roy. Soc.* A, **151**, 585.
Fuchs 1936*a* *Proc. Roy. Soc.* A, **153**, 622.
Fuchs 1936*b* *Proc. Roy. Soc.* A, **157**, 444.
Fuchs 1940 *Proc. Roy. Soc.* A, **176**, 214.
Mott & Jones 1936 *The theory of the properties of metals and alloys.* Oxford.
Moulton 1902 *Celestial mechanics.* Macmillan.
Peng 1941 *Proc. Roy. Soc.* A, **178**, 499.
Seitz 1935 *Phys. Rev.* **47**, 559.
Slater 1934 *Rev. Mod. Phys.* **6**, 209.
Wigner 1934 *Phys. Rev.* **46**, 1002.
Wigner & Seitz 1933 *Phys. Rev.* **43**, 804.
Wigner & Seitz 1934 *Phys. Rev.* **46**, 509.

Diffuse scattering of X-rays by crystals

II. Detailed calculation of the surfaces of isodiffusion for the (002), (112), (222) and (110) reflexions of sodium single crystals

BY H. A. JAHN

Davy-Faraday Laboratory, Royal Institution

(*Communicated by Sir William Bragg.—Received* 20 *November* 1941)

In connexion with the experimental work of Lonsdale & Smith on the diffuse scattering of X-rays by single crystals of sodium, the results obtained in previous publications have been completed by a detailed calculation of the surfaces of isodiffusion for the (002), (112), (222) and (110) reflexions of sodium single crystals. The surfaces are plotted in stereographic projection, and it is shown in a forthcoming publication of Lonsdale & Smith that these theoretical predictions are in excellent agreement with experiment. The Faxén-Waller theory of the diffuse scattering process has thus received decisive experimental confirmation.

INTRODUCTION

In a previous paper (Jahn 1942, referred to in the following as I) a general expression was found for the shapes of the surfaces of isodiffusion surrounding the various reciprocal lattice points of a cubic crystal. Simple expressions were found also for the relative intercepts of these surfaces along the major crystallographic directions. These particular values were evaluated for sodium single crystals and shown to indicate a large departure from spherical isotropy for the surfaces of isodiffusion and, moreover, a marked difference in the shape of the surfaces belonging to different crystal planes. This was illustrated also by the calculation of a typical plane section of the surfaces (see also Jahn 1941). In the present paper the surfaces of isodiffusion for the (002), (112), (222) and (110) reflexions have been evaluated in their entirety and plotted in stereographic projection. The marked departure from spherical isotropy and the large difference between surfaces belonging to different crystal planes is clearly illustrated (see figures 1 *a, b, c, d* below). In a forthcoming publication, Lonsdale & Smith (1942) show that these theoretical predictions are in remarkably close agreement with the results of their experimental investigation of the diffuse scattering of X-rays by single crystals of sodium.

1. SIMPLE GEOMETRY OF THE ADDITIONAL REFLEXIONS FOR SODIUM

The main features of the surfaces of isodiffusion for sodium calculated in this paper may be predicted in the following simple manner. As pointed out in my previous paper (I), the appearance of peaks in the surfaces of isodiffusion for sodium are due to the smallness of the fundamental elastic constant $c_{11}-c_{12}$ relative to the other elastic constants. It was shown in (I) that the elastic waves in a cubic crystal propagated along any one of the twelve equivalent digonal directions consist of one pure longitudinal and two pure transverse waves, each of different frequency. It was found that the square of the frequency of one of these two transverse waves was proportional to the fundamental constant $c_{11}-c_{12}$ alone. Consequently, these particular transverse waves will have, relatively speaking, exceptionally large amplitudes. Now it was shown in (I) that the additional Faxén-Waller reflexions (diffuse scattering) due to a single elastic wave were determined in direction by the direction of propagation of the elastic wave according to the Faxén-Waller vector condition ($\mathbf{B} = \mathbf{b}_n + \mathbf{W}$, see (I)) and in intensity by the product of the square of the amplitude of the wave ($q^2 \sim kT/\nu_2$ for $h\nu \ll kT$) with the square of the cosine of the angle between the direction of polarization ($\mathbf{e}$) of the wave with the normal ($\mathbf{b}_n$) to the particular set of planes under consideration. It is easy to see from the above, therefore, that the major features of the isodiffusion surfaces for sodium will consist of peaks along the twelve digonal directions, having relative intensities determined by the purely geometrical factor $(\mathbf{b}_n, \mathbf{e})^2$, where $\mathbf{e}$ denotes the polarization vector of that particular transverse wave referred to above, for each of the twelve directions of propagation in question, and $\mathbf{b}_n$ the normal to the set of planes giving the corresponding Bragg reflexion.

In table 1 the values of this expression for six of these directions (those for the remaining six follow from the centre of symmetry of the surface of isodiffusion) are listed for the (002), (112), (222) and (110) reflexions. In the last column the values for the general reflexion with indices (h_1, h_2, h_3) are also given. Comparison of this table with the surfaces plotted in figures 1 *a, b, c, d* below show that the relative strengths of the peaks and the manner in which they appear or disappear for the various reflexions are determined in the main by this simple factor. For certain planes and certain directions $(\mathbf{b}_n, \mathbf{e})$ is zero, i.e. $\mathbf{e}$ is perpendicular to the normal $\mathbf{b}_n$ of the reflecting plane and thus lies in the plane itself. It is clear that no displacement of the atoms in the plane itself can affect the Bragg reflexions from that plane and the surfaces show no trace of peaks for these cases.

TABLE 1. VALUES OF $(\mathbf{b}_n, \mathbf{e})^2$ FOR $\mathbf{b}_n$ AND $\mathbf{e}$ NORMALIZED TO UNITY

w	$\mathbf{e}$ \ $\mathbf{b}_n$	(002)	(112)	(222)	(110)	$(h_1 h_2 h_3)$
(110)	$(\bar{1}10)$	0	0	0	0	$\frac{(h_1-h_2)^2}{2(h_1^2+h_2^2+h_3^2)}$
(101)	$(10\bar{1})$	$\frac{1}{2}$	$\frac{1}{12}$	0	$\frac{1}{4}$	$\frac{(h_1-h_3)^2}{2(h_1^2+h_2^2+h_3^2)}$
(011)	$(01\bar{1})$	$\frac{1}{2}$	$\frac{1}{12}$	0	$\frac{1}{4}$	$\frac{(h_2-h_3)^2}{2(h_1^2+h_2^2+h_3^2)}$
$(\bar{1}10)$	(110)	0	$\frac{1}{3}$	$\frac{2}{3}$	1	$\frac{(h_1+h_2)^2}{2(h_1^2+h_2^2+h_3^2)}$
$(10\bar{1})$	(101)	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{2}{3}$	$\frac{1}{4}$	$\frac{(h_1+h_3)^2}{2(h_1^2+h_2^2+h_3^2)}$
$(01\bar{1})$	(011)	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{2}{3}$	$\frac{1}{4}$	$\frac{(h_2+h_3)^2}{2(h_1^2+h_2^2+h_3^2)}$

2. SURFACES OF ISODIFFUSION IN STEREOGRAPHIC PROJECTION

In calculating the complete surface of isodiffusion use is made of the general formula given in (I) for the square of the radius vector r drawn from the reciprocal lattice point in the direction (lmn) to meet the surface. It was found

$$r^2(lmn) = \frac{A(LMN, lmn)}{\Delta(lmn)},$$

where

$$\Delta(lmn) = c_{11}c_{44}^2 + c_{44}(c_{11}+c_{12})(c_{11}-c_{12}-2c_{44})(l^2m^2+l^2n^2+m^2n^2) + (c_{11}+2c_{12}+c_{44})(c_{11}-c_{12}-2c_{44})^2 l^2m^2n^2,$$

and $A(LMN, lmn)$

$$
\begin{aligned}
&= L^2\{c_{44}^2 + c_{44}(c_{11}-c_{44})(m^2+n^2) + (c_{11}+c_{12})(c_{11}-c_{12}-2c_{44})m^2n^2\} \\
&\quad + M^2\{c_{44}^2 + c_{44}(c_{11}-c_{44})(l^2+n^2) + (c_{11}+c_{12})(c_{11}-c_{12}-2c_{44})l^2n^2\} \\
&\quad + N^2\{c_{44}^2 + c_{44}(c_{11}-c_{44})(l^2+m^2) + (c_{11}+c_{12})(c_{11}-c_{12}-2c_{44})l^2m^2\} \\
&\quad - 2LM(c_{12}+c_{44})\{c_{44} + (c_{11}-c_{12}-2c_{44})n^2\}lm \\
&\quad - 2LN(c_{12}+c_{44})\{c_{44} + (c_{11}-c_{12}-2c_{44})m^2\}ln \\
&\quad - 2MN(c_{12}+c_{44})\{c_{44} + (c_{11}-c_{12}-2c_{44})l^2\}mn,
\end{aligned}
$$

where (LMN) are the direction cosines of the reciprocal lattice point in question (i.e. the direction cosines of the normal to the particular set of crystal planes under consideration), and c_{11}, c_{12}, c_{44} are the three elastic constants of the cubic crystal in standard notation.

The stereographic projection is introduced by putting

$$
l = \frac{2R}{1+R^2}\cos\phi, \qquad m = \frac{2R}{1+R^2}\sin\phi, \qquad n = \frac{1-R^2}{1+R^2},
$$

i.e. the values of the radius vector are regarded as a function plotted on a unit sphere centred at the reciprocal lattice point and this function projected from the south pole ((00$\bar{1}$) direction) of the sphere on to the equatorial plane (containing the (100) and (010) directions), ϕ being the angle measured from the 100 direction towards the 010 direction in the projection and R the radial distance from the origin of the projection. To facilitate comparison with the experimental work of Lonsdale & Smith, in which the crystal is usually oriented with the 1$\bar{1}$0 direction vertical, this direction has been chosen vertical in the stereographic projections printed in this paper.

From the above formulae the radius vector r was calculated for 0·1 intervals in the value of R and 5° intervals in the value of ϕ for the (002), (112), (222) and (110) reflexions of sodium as shown in tables 2A, B, C and D below (the limits of ϕ can be restricted in the manner shown for each case

TABLE 2A. VALUES OF r FOR SODIUM (002) REFLEXION
(FOR $R = 0$, $r = 1·0$)

R / ϕ	0·1	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9	1·0
0	1·06	1·28	1·71	2·10	1·90	1·51	1·26	1·13	1·06	1·0
5	1·06	1·28	1·70	2·08	1·88	1·50	1·26	1·13	1·06	1·0
10	1·06	1·28	1·67	2·01	1·84	1·49	1·25	1·12	1·06	1·0
15	1·06	1·27	1·64	1·92	1·77	1·46	1·25	1·12	1·06	1·0
20	1·06	1·27	1·59	1·82	1·69	1·43	1·23	1·12	1·06	1·0
25	1·06	1·27	1·55	1·72	1·60	1·39	1·22	1·11	1·06	1·0
30	1·06	1·26	1·51	1·63	1·52	1·34	1·20	1·11	1·06	1·0
35	1·06	1·25	1·48	1·57	1·46	1·30	1·17	1·10	1·06	1·0
40	1·06	1·25	1·46	1·53	1·42	1·27	1·16	1·09	1·06	1·0
45	1·06	1·24	1·46	1·52	1·40	1·26	1·15	1·09	1·06	1·0

TABLE 2B. VALUES OF r FOR SODIUM (112) REFLEXION
(FOR $R = 0$, $r = 1{\cdot}02$)

R ϕ	0·1	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9	1·0
45	0·885	0·828	0·832	0·852	0·850	0·845	0·853	0·877	0·911	0·957
40	0·886	0·830	0·835	0·857	0·856	0·851	0·858	0·879	0·913	0·958
35	0·888	0·834	0·844	0·873	0·875	0·867	0·869	0·886	0·916	0·960
30	0·891	0·841	0·859	0·900	0·903	0·888	0·883	0·894	0·921	0·964
25	0·896	0·850	0·879	0·936	0·938	0·912	0·897	0·902	0·926	0·969
20	0·901	0·863	0·904	0·978	0·977	0·934	0·909	0·910	0·934	0·976
15	0·908	0·877	0·931	1·02	1·01	0·953	0·920	0·919	0·942	0·986
10	0·916	0·894	0·960	1·07	1·05	0·969	0·930	0·929	0·953	0·999
5	0·926	0·912	0·989	1·11	1·07	0·981	0·941	0·942	0·969	1·02
0	0·936	0·932	1·02	1·13	1·08	0·992	0·954	0·958	0·988	1·04
− 5	0·947	0·955	1·04	1·15	1·09	1·00	0·971	0·981	1·02	1·07
− 10	0·959	0·979	1·07	1·16	1·10	1·02	0·995	1·01	1·05	1·11
− 15	0·973	1·00	1·10	1·17	1·11	1·04	1·03	1·06	1·10	1·16
− 20	0·986	1·03	1·13	1·18	1·13	1·08	1·08	1·12	1·17	1·24
− 25	1·00	1·06	1·17	1·21	1·16	1·13	1·15	1·20	1·27	1·34
− 30	1·02	1·09	1·21	1·25	1·21	1·20	1·24	1·32	1·40	1·47
− 35	1·03	1·13	1·26	1·31	1·28	1·28	1·35	1·46	1·57	1·64
− 40	1·05	1·16	1·31	1·38	1·36	1·38	1·47	1·60	1·73	1·80
− 45	1·06	1·20	1·38	1·47	1·46	1·48	1·57	1·70	1·83	1·87
− 50	1·08	1·24	1·46	1·58	1·57	1·57	1·63	1·72	1·79	1·80
− 55	1·10	1·28	1·54	1·70	1·69	1·65	1·65	1·67	1·67	1·64
− 60	1·11	1·32	1·63	1·83	1·82	1·72	1·64	1·58	1·53	1·47
− 65	1·13	1·36	1·72	1·98	1·94	1·76	1·61	1·50	1·41	1·34
− 70	1·14	1·40	1·81	2·13	2·05	1·79	1·57	1·42	1·31	1·24
− 75	1·15	1·44	1·89	2·28	2·15	1·80	1·53	1·36	1·24	1·16
− 80	1·17	1·47	1·97	2·41	2·22	1·80	1·50	1·31	1·19	1·11
− 85	1·18	1·49	2·02	2·50	2·26	1·79	1·47	1·27	1·15	1·07
− 90	1·19	1·51	2·05	2·54	2·27	1·77	1·44	1·24	1·11	1·04
− 95	1·20	1·52	2·06	2·51	2·24	1·74	1·41	1·21	1·09	1·02
− 100	1·21	1·53	2·04	2·44	2·17	1·71	1·38	1·19	1·07	0·999
− 105	1·21	1·54	2·01	2·33	2·08	1·66	1·36	1·17	1·06	0·986
− 110	1·22	1·54	1·96	2·21	1·98	1·61	1·33	1·16	1·04	0·976
− 115	1·22	1·54	1·92	2·09	1·88	1·56	1·31	1·14	1·03	0·969
− 120	1·23	1·54	1·88	1·99	1·79	1·51	1·29	1·13	1·03	0·964
− 125	1·23	1·53	1·84	1·92	1·72	1·46	1·26	1·12	1·02	0·960
− 130	1·23	1·53	1·82	1·87	1·67	1·43	1·24	1·11	1·02	0·958
− 135	1·23	1·53	1·82	1·86	1·66	1·42	1·24	1·11	1·02	0·957

TABLE 2C. VALUES OF r FOR SODIUM (222) REFLEXION
(FOR $R = 0$, $r = 1{\cdot}03$)

ϕ \ R	0·1	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9	1·0
45	0·90	0·81	0·75	0·71	0·70	0·70	0·73	0·76	0·81	0·86
40	0·90	0·81	0·76	0·72	0·70	0·71	0·73	0·76	0·81	0·86
35	0·90	0·81	0·76	0·72	0·71	0·71	0·74	0·77	0·81	0·87
30	0·90	0·82	0·77	0·73	0·72	0·73	0·75	0·78	0·82	0·87
25	0·90	0·83	0·77	0·74	0·73	0·74	0·76	0·79	0·83	0·89
20	0·91	0·84	0·79	0·76	0·75	0·76	0·78	0·81	0·85	0·90
15	0·92	0·85	0·80	0·78	0·77	0·78	0·80	0·83	0·87	0·92
10	0·92	0·86	0·82	0·80	0·80	0·81	0·83	0·86	0·90	0·95
5	0·93	0·87	0·84	0·83	0·83	0·84	0·86	0·89	0·93	0·99
0	0·94	0·89	0·87	0·86	0·86	0·88	0·90	0·93	0·98	1·03
−5	0·95	0·91	0·90	0·89	0·90	0·92	0·95	0·99	1·03	1·09
−10	0·96	0·93	0·93	0·94	0·95	0·98	1·02	1·06	1·11	1·17
−15	0·98	0·96	0·97	0·98	1·01	1·06	1·10	1·15	1·21	1·27
−20	0·99	0·99	1·01	1·04	1·08	1·14	1·21	1·28	1·34	1·40
−25	1·00	1·02	1·06	1·10	1·15	1·24	1·34	1·43	1·51	1·58
−30	1·02	1·05	1·11	1·16	1·24	1·35	1·49	1·62	1·73	1·80
−35	1·03	1·09	1·17	1·24	1·32	1·46	1·65	1·84	1·99	2·07
−40	1·05	1·12	1·23	1·32	1·42	1·57	1·79	2·04	2·23	2·32
−45	1·06	1·16	1·30	1·42	1·52	1·67	1·89	2·15	2·36	2·43
−50	1·08	1·20	1·38	1·53	1·62	1·75	1·93	2·13	2·28	2·32
−55	1·10	1·25	1·46	1·64	1·73	1·80	1·90	2·01	2·07	2·07
−60	1·11	1·29	1·55	1·77	1·83	1·83	1·84	1·85	1·84	1·80
−65	1·13	1·33	1·64	1·91	1·93	1·84	1·75	1·69	1·63	1·58
−70	1·14	1·37	1·73	2·05	2·03	1·83	1·67	1·55	1·47	1·40
−75	1·16	1·40	1·82	2·19	2·10	1·81	1·58	1·44	1·34	1·27
−80	1·17	1·44	1·89	2·31	2·16	1·78	1·51	1·35	1·24	1·17
−85	1·18	1·46	1·95	2·39	2·18	1·75	1·45	1·28	1·16	1·09
−90	1·19	1·49	1·98	2·42	2·17	1·71	1·40	1·22	1·10	1·03
−95	1·20	1·50	1·99	2·37	2·13	1·66	1·36	1·17	1·06	0·99
−100	1·21	1·51	1·98	2·33	2·06	1·62	1·32	1·13	1·02	0·95
−105	1·22	1·52	1·95	2·23	1·98	1·57	1·28	1·10	0·99	0·92
−110	1·23	1·52	1·91	2·12	1·88	1·52	1·25	1·08	0·97	0·90
−115	1·23	1·52	1·87	2·01	1·79	1·47	1·22	1·06	0·95	0·89
−120	1·23	1·52	1·84	1·92	1·71	1·42	1·20	1·04	0·94	0·87
−125	1·24	1·52	1·81	1·86	1·64	1·38	1·17	1·03	0·93	0·87
−130	1·24	1·52	1·79	1·81	1·60	1·35	1·16	1·02	0·93	0·86
−135	1·24	1·52	1·78	1·80	1·58	1·34	1·15	1·02	0·92	0·86

TABLE 2D. VALUES OF r FOR SODIUM (110) REFLEXION
(FOR $R = 0$, $r = 1{\cdot}05$)

R / ϕ	0·1	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9	1·0
45	1·09	1·21	1·32	1·29	1·12	0·962	0·853	0·788	0·756	0·748
40	1·09	1·21	1·33	1·30	1·13	0·969	0·856	0·792	0·759	0·751
35	1·09	1·21	1·33	1·32	1·15	0·982	0·865	0·800	0·767	0·758
30	1·09	1·21	1·35	1·36	1·19	1·01	0·881	0·813	0·781	0·772
25	1·09	1·21	1·37	1·41	1·23	1·03	0·904	0·833	0·801	0·792
20	1·09	1·21	1·40	1·47	1·29	1·07	0·933	0·861	0·828	0·819
15	1·09	1·21	1·41	1·53	1·35	1·11	0·969	0·896	0·863	0·854
10	1·08	1·20	1·43	1·59	1·41	1·16	1·01	0·940	0·907	0·898
5	1·08	1·20	1·44	1·64	1·47	1·22	1·07	0·995	0·962	0·953
0	1·08	1·20	1·44	1·66	1·52	1·28	1·14	1·06	1·03	1·02
−5	1·08	1·19	1·42	1·66	1·55	1·34	1·21	1·15	1·12	1·11
−10	1·07	1·17	1·39	1·62	1·58	1·41	1·31	1·25	1·23	1·22
−15	1·07	1·17	1·36	1·58	1·59	1·49	1·42	1·39	1·36	1·36
−20	1·07	1·15	1·32	1·53	1·59	1·57	1·56	1·55	1·55	1·55
−25	1·07	1·14	1·29	1·47	1·59	1·66	1·70	1·75	1·77	1·78
−30	1·07	1·14	1·26	1·43	1·58	1·73	1·86	1·98	2·05	2·08
−35	1·06	1·13	1·24	1·39	1·58	1·79	2·01	2·23	2·37	2·42
−40	1·06	1·12	1·22	1·37	1·54	1·83	2·12	2·44	2·66	2·75
−45	1·06	1·12	1·22	1·37	1·54	1·84	2·17	2·51	2·79	2·89

owing to the symmetry of the surface; in the case of the (222) surface, which has a six-fold axis of symmetry, the calculation could have been simplified by projection from the (111) direction, but we have preferred to keep the projection the same in all cases to facilitate comparison). The same relative values of the elastic constants were used as in the previous paper (viz. $c_{12} = 0{\cdot}760\,c_{11}$, $c_{44} = 0{\cdot}914\,c_{11}$) and c_{11} put equal to unity, so that these tables can be regarded as an extension of table 2 in that paper.

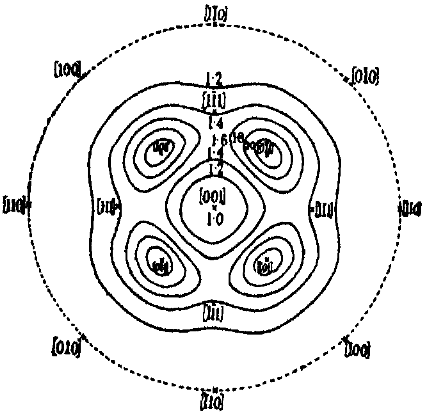


FIGURE 1a. Stereographic projection of surface of isodiffusion surrounding (002) reciprocal lattice point for sodium.

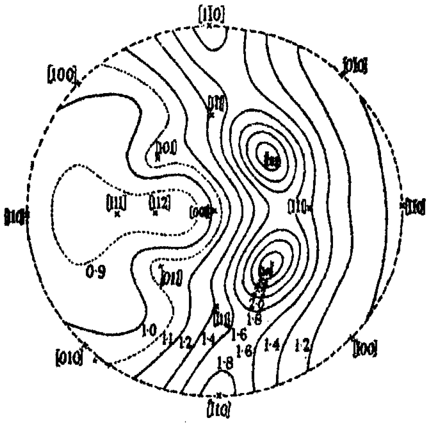


FIGURE 1b. Stereographic projection of surface of isodiffusion surrounding (112) reciprocal lattice point for sodium.

From these numerical values the stereographic projection of the contours of equal radius on the surface of isodiffusion were plotted and the results are shown in figures 1 *a*, *b*, *c*, *d*. A comparison with the experimental results will be given in the publication of Lonsdale & Smith referred to above.

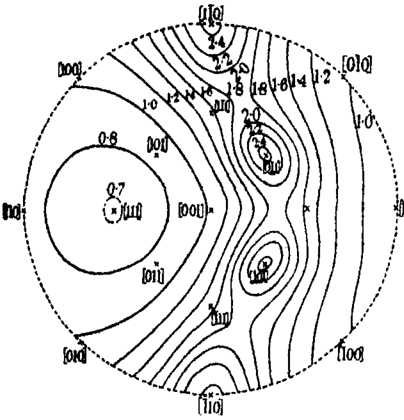


FIGURE 1*c*. Stereographic projection of surface of isodiffusion surrounding (222) reciprocal lattice point for sodium.

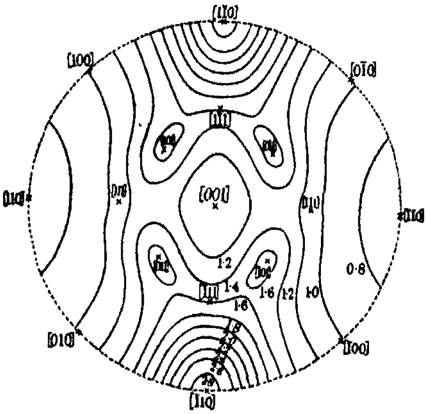


FIGURE 1*d*. Stereographic projection of surface of isodiffusion surrounding (110) reciprocal lattice point for sodium.

CONCLUSION

The agreement of these theoretical predictions with the experimental results of Lonsdale & Smith provides decisive experimental confirmation of the Faxén-Waller theory of the diffuse scattering process. I am convinced that intensity measurements made on the diffusely scattered X-rays from single crystals will lead eventually to an X-ray method of determining the elastic constants of crystals. Due correction would of course have to be made for the incoherent scattering and the accuracy needed may require improvements in the experimental technique, but it is clear that an important field of research has been opened up. The fact that, for elastically anisotropic single crystals, the reciprocal lattice points are surrounded by differently shaped regions of scattering power will clearly have an important bearing on the conclusions drawn about particle size and shape from the intensity distribution in Debye-Scherrer lines of polycrystalline material.

REFERENCES

Jahn, H. A. 1941 *Nature, Lond.*, **147**, 511.
Jahn, H. A. 1942 *Proc. Roy. Soc.* **179**, 320.
Lonsdale, K. & Smith, H. 1942 *Nature, Lond.*, **149**, 21.

The free streaming of gases in sloping galleries

By E. H. M. Georgeson

(*Communicated by G. I. Taylor, F.R.S.—Received* 26 *November* 1941)

A general equation is deduced for the free streaming of any gas in any sloping, rectangular, air-filled gallery. Values for the velocities and accelerations of freely streaming firedamp and hydrogen are calculated from this equation, and are compared with experimental results.

During the study of some problems of mine ventilation it appeared that when firedamp was allowed to stream freely up the sloping roof of a gallery, the front of the body of gas moved through the air for a considerable distance without measurable acceleration—a somewhat unexpected finding. This paper gives a simple analysis of the behaviour of such a streaming gas, and describes experiments designed to elucidate the apparent constancy of speed.

The experiments were carried out in a concrete gallery at the Buxton experimental station of the Safety in Mines Research Board. The gallery is 100 ft. long, 6 ft. high, 5 ft. wide and has a slope of 1 in 10. The gas used was piped through a meter to a flat box supported against the roof at the closed lower end of the gallery. This box filled the whole width of the gallery, and the gas issued from it as a shallow stream almost 5 ft. wide. The flow up the roof was indicated in the earlier experiments by a row of top-feed safety lamps hung as nearly flush with the roof as was possible and spaced 12 ft. 6 in. apart up the middle of the gallery. Later, feathers and thistledowns (*Cnicus eriophorus*), tethered by hairs and fine 'Durofix' (trinitrocellulose) fibres, proved to be more sensitive and adaptable indicators. They cause no perceptible disturbance and respond instantly to very small movements of gas or air.

In 107 experiments, during each of which firedamp poured continuously into the gallery at some measured rate of entry, the speed of streaming when measured over the stretch from 25 to 75 ft. from the inlet was constant within the rather large limits of experimental error. A summary of the speeds found with various rates of inflow appears in table 1. The points are shown in relation to the calculated velocities in figure 1, and the results of individual experiments are in the appendix.

The behaviour of a much lighter gas was also studied as it seemed likely to help in solving the problem. Table 2 summarizes the results of thirty-three experiments in which hydrogen streamed freely up the roof of the airway. A small acceleration was detected at the highest rate of inflow.

TABLE 1

number of experiments averaged	rate of entry of firedamp into gallery cu. ft./min.	range	mean speed of streaming ft./sec.
4	6·3	6·0– 6·9	0·87
3	7·9	7·8– 7·9	0·94
4	9·5	9·4– 9·6	0·99
37	10·0	9·8–10·3	1·07
7	10·9	10·7–11·2	1·14
2	12·1	12·0–12·3	1·08
9	15·2	14·8–16·0	1·16
4	18·8	18·0–19·2	1·29
37	20·0	19·3–21·1	1·31

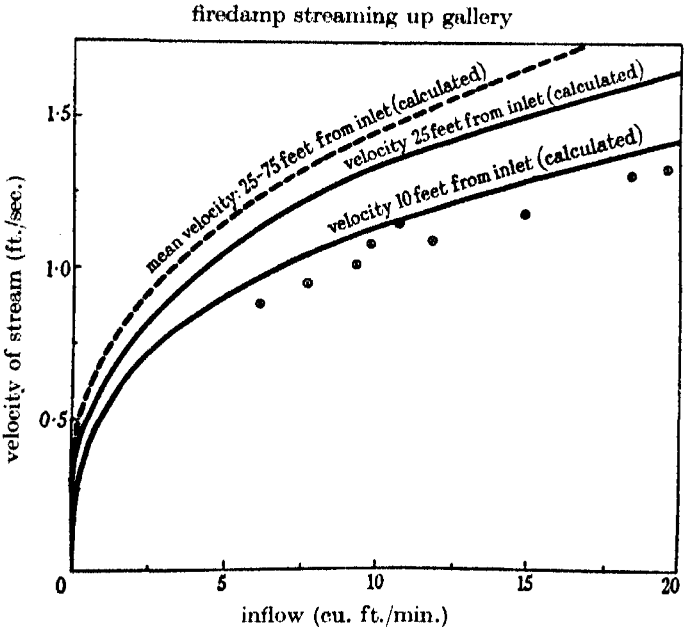


FIGURE 1

These points are shown in relation to the calculated velocities in figure 2, and the results of individual experiments are in the Appendix.

It is so usual in everyday life to think in terms of the slow motion of solid bodies in the light fluid air that the moving mass is often considered invariable. This is not quite true; because unless the solid body moves in a vacuum some part of the surrounding matter is bound to receive and to retain some part of its momentum and so, inevitably, the moving mass grows. The present work deals with the movement of a body of gas travelling through gas, and here the deeper interplay of the molecules facilitates the transfer of momentum from the moving body of gas to the ambient air.

TABLE 2

number of experiments averaged	rate of entry of firedamp into gallery cu. ft./min.	range	mean speed of streaming ft./sec.
2	1·1	1·1– 1·2	0·87
4	2·3	2·0– 2·5	1·09
2	4·3	4·2– 4·4	1·22
4	7·7	7·1– 8·3	1·28
2	9·3	9·3	1·60
12	10·0	9·4–10·9	1·45
1	12·0	12·0	1·57
6	20·9	20·2–21·9	1·57

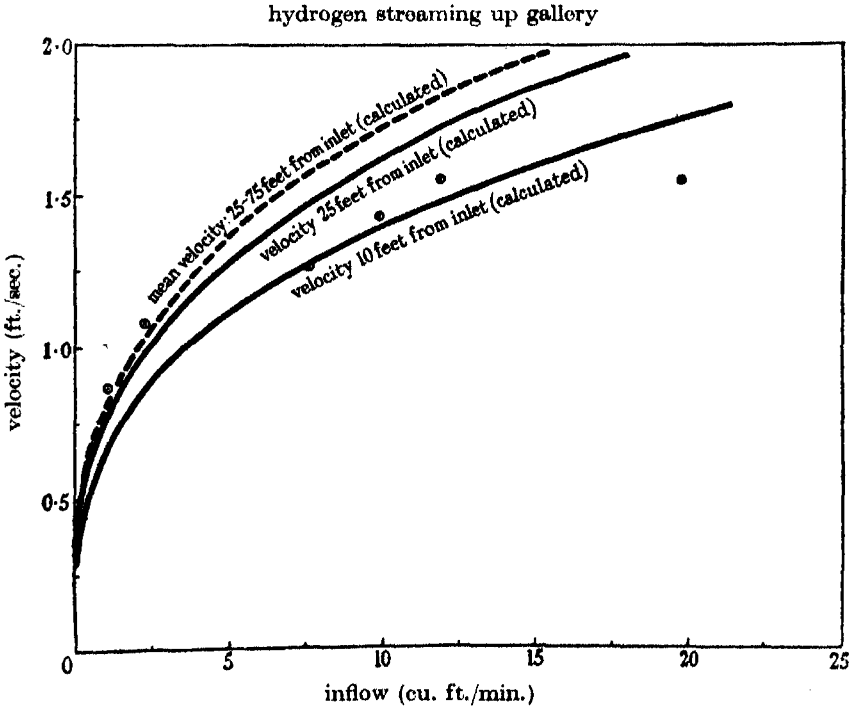


FIGURE 2

Further, the comparatively slight difference in density between the body of gas and the air magnifies the influence of the growth of the mass on the inhibition of acceleration.

Is there, as Helmholtz argued for somewhat similar circumstances, a sharp velocity discontinuity at some stratum a definite distance below the roof, or does the velocity fade away as the depth below the roof increases? Thistledown indicators hung at different heights in the gallery have proved conclusively that the velocity drops from about 1 ft./sec. to zero in a vertical distance of less than 1 in., and have shown further that the steadily moving stream deepens from 8 in. at the cross-section 25 ft. from the inlet to 14 in. 75 ft. from the inlet, and that the velocity of the stream is constant throughout its depth. Furthermore, although there is a deeper 'head' at the front, the depth of the body of the stream at these two points does not vary by more than ½ in. whether the rate of inflow is as low as 5 or as high as 20 cu. ft. of firedamp per minute. The depths at these points are the same, indeed, even when hydrogen is the streaming gas.

There is evidence (Richardson 1920; Taylor 1922) that the cloud of smoke from a chimney is sometimes a paraboloid whose vertex is at the chimney-top. This means that if a is the radius of any cross-section of the cloud distant s from the chimney-top, $a = Rs^{\frac{1}{2}}$. This leads to the conjecture that the depth (h) of the stream in the gallery may vary as $s^{\frac{1}{2}}$. From the measurements given above we have two pairs of values of h and s:

$$\tfrac{8}{12} = R\sqrt{(25)}, \quad \text{whence } R = 0{\cdot}1333,$$

$$\tfrac{14}{12} = R\sqrt{(75)}, \quad \text{whence } R = 0{\cdot}1347.$$

This is good agreement, and $h = 0{\cdot}134,\ s^{\frac{1}{2}}$ may be taken to be the depth of the stream at any point. The outline of the stream in the gallery is shown in figure 3.

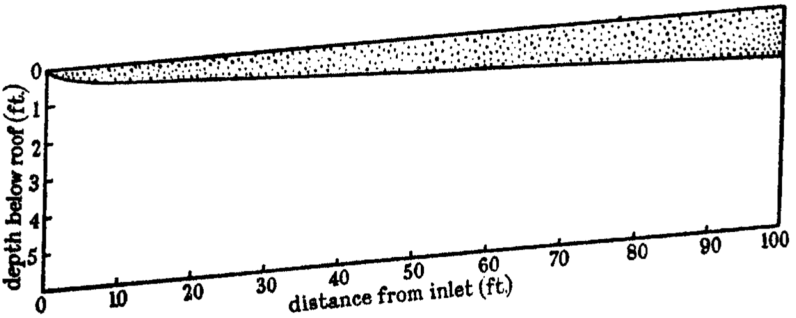


FIGURE 3

For steady motion up the roof of the gallery the equation of motion may be written

$$\rho_{\text{air}}\, v\frac{\partial v}{\partial s} = -\frac{\partial p}{\partial s} + X, \tag{1}$$

where ρ_{air} is the density of air (lb./cu.ft.),

v is the velocity at any point distant s ft. from the inlet (ft./sec.),

p is the pressure (lb./sq. ft.), and

X is the extraneous force acting in the direction of the length of the gallery (lb.).

So long as the depth of the gallery is great compared with the depth of the gas stream $\partial p/\partial s$ will be small compared with X, and so, in these circumstances, we may write

$$\rho_{\text{air}}\frac{\partial}{\partial s}(\tfrac{1}{2}v^2) = X. \tag{2}$$

X is composed of two parts, one that due to the component of gravity which acts in the direction of the length of the gallery, the other that due to the transfer of momentum from the moving stream of gas to the still air beneath. The first part may be written

$$\frac{Qg\sin\alpha}{60B}(\rho_{\text{air}}-\rho_{\text{gas}})\frac{1}{hv},$$

where Q is the rate of inflow of gas into the gallery (cu. ft./min.),

ρ_{gas} is the density of the inflowing gas (lb./cu. ft.),

h is the depth of the stream at any point distant s ft. from the inlet (ft.),

B is the breadth of the gallery (ft.), and

α is its angle of inclination from the horizontal.

The transfer of momentum due to the inclusion of a mass of air of $m\,\delta s$ per sec. into a section of the stream δs long, moving with velocity v is $mv\,\delta s$ and the drag per unit volume is mv/h.

The change in depth, i.e. $(dh/ds)\,\delta s$, is due partly to the inclusion of extra air and partly to the alteration in the speed of the stream. Thus the equation of continuity applied to the volume bounded by two sections of the stream distant δs apart and the base of the stream is

$$-v\frac{dh}{ds}\delta s - h\frac{dv}{ds}\delta s + \frac{m}{\rho_{\text{air}}}\delta s = 0. \tag{3}$$

The two parts of X having been evaluated, equation (2) may be rewritten thus:

$$\rho_{\text{air}}\frac{d}{ds}(\tfrac{1}{2}v^2) = \frac{Qg\sin\alpha}{60B}(\rho_{\text{air}}-\rho_{\text{gas}})\frac{1}{hv} - \frac{mv}{h}. \tag{4}$$

From (3)

$$m = h\rho_{\text{air}}\left\{\frac{dv}{ds}+\frac{v}{h}\frac{dh}{ds}\right\}. \tag{5}$$

Hence (4) becomes

$$\rho_{\text{air}}\frac{d}{ds}v^2 = \frac{Qg\sin\alpha}{60B}\frac{(\rho_{\text{air}}-\rho_{\text{gas}})}{hv} - \rho_{\text{air}}\frac{v^2}{h}\frac{dh}{ds}$$

or

$$\tfrac{2}{3}h\frac{d}{ds}v^3 + v^3\frac{dh}{ds} = \frac{Qg\sin\alpha}{60B}\left(\frac{\rho_{\text{air}}-\rho_{\text{gas}}}{\rho_{\text{air}}}\right) = \mathrm{A},$$

whence
$$h^{3/2}\frac{d}{ds}v^3+\tfrac{3}{2}h^{1/2}v^3\frac{dh}{ds}=\tfrac{3}{2}h^{1/2}\mathrm{A}.$$

or
$$\frac{d}{ds}(h^{3/2}v^3)=\tfrac{3}{2}h^{1/2}\mathrm{A},$$

whence
$$h^{3/2}v^3=\tfrac{3}{2}\mathrm{A}\int h^{1/2}\,ds+\text{const.}\tag{6}$$

This is a general expression, including all possible relationships between h and s; but since it has been found experimentally that $h = Rs^{1/2}$,

$$\int h^{1/2}\,ds = R^{1/2}\int s^{1/4}\,ds = \tfrac{4}{5}R^{1/2}s^{5/4}.$$

Hence, then, $h^{3/2}v^3 = \tfrac{3}{2}\mathrm{A}(\tfrac{4}{5})\,R^{1/2}s^{5/4}+\text{const.}$

or
$$R^{3/2}s^{3/4}v^3=\tfrac{6}{5}\mathrm{A}\,R^{1/2}s^{5/4}+\text{const.},$$

whence
$$v^3=\tfrac{6}{5}\mathrm{A}s^{1/2}R^{-1}+\text{const.}\,(s^{-3/4}).\tag{7}$$

Since the initial conditions are not consistent with an infinite velocity at the point of inflow, the constant in (7) is zero.

Hence, then
$$v=\left\{\frac{6}{5}\left(\frac{Qg\sin\alpha}{60B}\right)\left(\frac{\rho_{\text{air}}-\rho_{\text{gas}}}{\rho_{\text{air}}}\right)\frac{1}{R}\right\}^{1/3}s^{1/6}$$

or
$$v=0{\cdot}271\left\{\frac{Qg\sin\alpha}{BR}\left(\frac{\rho_{\text{air}}-\rho_{\text{gas}}}{\rho_{\text{air}}}\right)\right\}^{1/3}s^{1/6},\tag{8}$$

which may perhaps be called the general equation for the streaming of gas in an air-filled, sloping, rectangular gallery. This equation takes into account only the balance between the buoyancy of the gas stream and the momentum interchange between it and the air beneath: viscosity and any skin friction between the stream and the roof of the gallery are neglected.

From this, when the values for the density of firedamp (0·045 lb./cu. ft.) and of air (0·080 lb./cu. ft.) and for the inclination and breadth of the Buxton gallery are inserted:

$$v=0{\cdot}347Q^{1/3}s^{1/6},\tag{9}$$

whence
$$\frac{dv}{dt}=0{\cdot}0578Q^{1/3}s^{-5/6}v$$
$$=0{\cdot}020Q^{2/3}s^{-2/3}.\tag{10}$$

The mean velocity from $s = s_0$ to $s = s_1$ may be written

$$\bar{v}_{s_0s_1}=\frac{0{\cdot}347Q^{1/3}\int_{s_0}^{s_1}s^{1/6}\,ds}{\int_{s_0}^{s_1}ds},$$

whence
$$\bar{v}_{s_0 s_1} = 0{\cdot}297Q^{\frac{1}{3}}\frac{s_1^{\frac{7}{6}} - s_0^{\frac{7}{6}}}{s_1 - s_0}. \tag{11}$$

The time taken to traverse any distance may be written

$$t_{s_0 s_1} = \frac{(s_1 - s_0)^2}{0{\cdot}298Q^{\frac{1}{3}}(s_1^{\frac{7}{6}} - s_0^{\frac{7}{6}})}. \tag{12}$$

Figure 1 has been drawn from equations (9) and (11) to show velocities in the gallery calculated for various rates of inflow of firedamp. The full lines show the velocities calculated for points distant 10, 25 and 50 ft. from the inlet. The dotted line shows the calculated mean velocities for the length 25–75 ft. from the inlet. The circles are the experimental points taken from table 1.

The mean percentage of the inflowing gas in any cross-section of the stream distant s ft. from the inlet is

$$\frac{Q}{60B\bar{v}_{s_0 s_1}}\,\frac{100}{Rs_1^{\frac{1}{2}}},$$

which simplifies to
$$8{\cdot}35Q^{\frac{2}{3}}s^{-\frac{2}{3}}.^*$$

For hydrogen (density 0·007 lb./cu. ft.) the corresponding equations are

$$v = 0{\cdot}443Q^{\frac{1}{3}}s^{\frac{1}{6}}, \tag{9a}$$

$$\frac{dv}{dt} = 0{\cdot}0327Q^{\frac{2}{3}}s^{-\frac{2}{3}}, \tag{10a}$$

$$\bar{v}_{s_0 s_1} = 0{\cdot}380Q^{\frac{1}{3}}\frac{s_1^{\frac{7}{6}} - s_0^{\frac{7}{6}}}{s_1 - s_0}, \tag{11a}$$

and
$$t_{s_0 s_1} = \frac{(s_1 - s_0)^2}{0{\cdot}380Q^{\frac{1}{3}}(s_1^{\frac{7}{6}} - s_0^{\frac{7}{6}})}. \tag{12a}$$

Figure 2 is similar to figure 1, hydrogen instead of firedamp being the streaming gas.

The mean percentage of hydrogen in any cross-section of the stream in the airway is $6{\cdot}54Q^{\frac{2}{3}}s^{-\frac{2}{3}}$.

---

* When safety lamps were used as indicators the one distant 37·5 ft. from the inlet was extinguished by each 20 cu. ft./min. stream, but the lamp 50 ft. from the inlet was never extinguished. The calculated percentages of firedamp present are tabulated below:

Q (cu. ft./min.)	distance from inlet (ft.)	calculated mean percentage of firedamp in stream
20	37·5	5·5
	50	4·5
	75	1·8

It is known that such lamps are extinguished when put into an atmosphere of air containing not less than from 5 to 6 % of methane.

Table 3 shows the calculated mean speeds of streaming in different sections of the gallery for inflow rates of 10 cu. ft./min. of firedamp, and of 9·5 and 20 cu. ft./min. of hydrogen. It also shows the corresponding calculated and observed times taken in traversing the sections.

TABLE 3

distance from inlet of section concerned (ft.)	firedamp inflow 10 cu. ft./min. time to traverse section (sec.)		hydrogen inflow 9·5 cu. ft./min. time to traverse section (sec.)		hydrogen inflow 20 cu. ft./min. time to traverse section (sec.)	
	calc.	obs.	calc.	obs.	calc.	obs.
0 – 12·5	12·2	18·0	—	—	7·6	—
0 – 20	—	—	15·1	21·2	—	17·5
12·5– 25	10·3	11·5	—	—	6·4	—
20 – 45	—	—	15·0	17·6	—	—
25 – 37·5	9·4	12·5	—	—	5·9	—
37·5– 50	8·9	11·0	—	—	5·5	17·3
45 – 70	—	—	13·6	15·6	—	—
50 – 62·5	8·5	13·0	—	—	5·3	—
62·5– 75	8·3	10·5	—	—	5·2	15·2
70 – 95	—	—	12·8	15·6	—	—
75 – 87·5	8·0	12·0	—	—	5·0	—
87·5–100	7·85	—	—	—	4·9	—
25 – 75	35·1	47·0	—	—	21·9	32·5
20 – 95	—	—	41·5	48·8	—	—

It may seem strange that throughout this paper, although R, the constant in the equation for the depth of the moving gas stream, is the only arbitrary parameter introduced, there has been no mention of diffusion, viscosity, or skin friction against presumably the airway. No doubt all these do retard the stream. Presumably they account for the fact that the speeds calculated from the simple equation are about one-fifth higher than the observed speeds.

That R remains the same for gases as dissimilar as firedamp and hydrogen suggests that it may depend only upon the viscosity of air and upon the degree of inclination of the gallery. Dimensionally, the equation $h = Rs^{\frac{1}{2}}$ should be of the form $h = \text{constant}\, s^{\frac{1}{2}}Q^{\frac{1}{2}}$. It is likely that the influence of $Q^{\frac{1}{2}}$ is masked by the narrow range of inflows considered.

The writer thanks Professor G. I. Taylor for constructive criticisms, and the Safety in Mines Research Board for permission to publish this paper.

REFERENCES

Richardson, L. F. 1920 *Phil. Trans.* A, **221**, 1.
Taylor, G. I. 1922 *Proc. Lond. Math. Soc.* (2), **20**, 196.

APPENDIX

TABLE 1. FIREDAMP

rate of entry of firedamp into gallery cu. ft./min.	speed of streaming ft./sec. (25–75 ft. from inlet)
6·0	0·82
6·1	0·90
6·3	0·89
6·9	0·87
7·8	0·95
7·9	1·00, 0·87
9·4	1·08, 0·89
9·6	1·02, 0·98
9·8	1·07
9·9	1·12, 1·12
10·0	1·17, 0·90, 1·01, 0·92, 1·08, 1·18, 1·25, 1·48, 0·93, 0·93, 1·05, 1·08, 1·00, 1·18, 1·13, 1·18, 1·13, 1·18, 1·13, 1·32, 1·00, 1·13, 0·97, 0·97, 0·97, 1·05, 1·05, 1·08
10·2	1·21, 1·17, 1·14, 1·12
10·3	1·05, 1·08
10·7	0·98, 1·00
10·8	1·13, 1·23
10·9	1·25
11·0	1·25
11·2	1·13
12·0	1·13
12·3	1·02
14·8	1·08
15·0	0·97, 1·18, 1·18, 1·47
15·1	1·13
15·2	1·17
15·4	1·17
16·0	1·07
18·0	1·28
19·0	1·35, 1·33
19·2	1·22
19·3	1·33
19·4	1·43, 1·33
19·8	1·25, 1·38, 1·43
19·9	1·58, 1·38
2·00	1·18, 1·57, 1·18, 1·00, 1·18, 1·25, 1·25, 1·37, 1·33, 1·25, 1·18, 1·47, 1·18, 1·23, 1·35, 1·35, 1·57, 1·47, 1·47, 1·12, 1·14, 1·23, 1·10, 1·25, 1·35
20·1	1·23
20·2	1·28
20·6	1·35
21·1	1·22

TABLE 2. HYDROGEN

rate of entry of hydrogen into airway cu. ft./min.	speeds of streaming (ft./sec.)			
	20–45 ft. from inlet	25–75 ft. from inlet	45–70 ft. from inlet	70–95 ft. from inlet
1·1	0·8	—	0·9	1·1
1·2	0·5	—	0·9	1·0
2·0	0·9	—	1·1	1·0
2·2	1·0	—	1·2	1·4
2·5	1·0	—	1·2	1·2
2·5	1·0	—	1·1	1·2
4·2	1·1	—	1·3	1·4
4·4	1·1	—	1·2	1·3
7·1	—	1·0	—	—
7·5	—	1·2	—	—
7·9	1·5	—	1·4	1·6
8·3	1·1	—	1·2	1·2
9·3	1·4	—	1·6	1·6
9·3	1·5	—	1·8	1·7
9·4	—	1·1	—	—
9·6	—	1·0	—	—
9·6	—	1·3	—	—
9·6	1·3	—	1·4	1·5
9·8	—	1·2	—	—
9·8	—	1·1	—	—
9·9	—	1·5	—	—
10·1	1·5	—	1·5	1·6
10·1	1·6	—	1·8	1·9
10·3	—	1·6	—	—
10·6	—	1·8	—	—
10·9	—	1·4	—	—
12·0	1·4	—	1·6	1·8
20·2	—	1·6	—	—
20·4	—	1·6	—	—
20·6	—	1·5	—	—
20·7	—	1·7	—	—
21·6	—	1·5	—	—
21·9	—	1·6	—	—

INDEX TO VOLUME 180 (A)

Aeolotropic plates, stress systems (Green), 173.
Autelectronic emission distribution from single crystal metal points (Benjamin & Jenkins), 225.

Bakerian Lecture (Dirac), 1.
Benjamin, M. & Jenkins, R. O. The distribution of autelectronic emission from single crystal metal points, 11, 225.
Bickley, W. G. The effect of a free surface on compressional (shock) waves, 209.
Born, M. Effect of thermal vibrations on the scattering of X-rays, III, 397.
Bunn, C. W. Molecular structure and rubber-like elasticity, 40, 67, 82.

Chang, T. S. & Ho, C. C. Arrangements with given number of neighbours, 345.
Clews, C. J. B. The structure of polychloroprene, 100.
Cobalt, imperfections in structure, I (Edwards & Lipson), 268. II (Wilson), 277.
Compressional (shock) waves, effect of a free surface (Bickley), 209.
Cosmic radiation, the 'second maximum' of the shower transition curve (George, Jánossy & McCaig), 219.
Crystal theory of metals, elastic constants (Fuchs & Peng), 451.

Dirac, P. A. M. The physical interpretation of quantum mechanics (Bakerian Lecture), 1.
Disintegrations produced by deuterons in heavy elements, cross-section measurements (Krishnan & Nahum), 321; excitation function measurements (Krishnan & Nahum), 333.

Edwards, O. S. & Lipson, H. Imperfections in the structure of cobalt, I, 268.
Elastic crystals, theory of finite deformations (Fürth), 285.

Free radical chain reactions, detection and inhibition (Smith & Hinshelwood), 237.
Fuchs, K. & Peng, H. W. Crystal theory of metals: calculation of the elastic constants, 451.
Fürth, R. On the theory of finite deformations of elastic crystals, 285.

Gases, free streaming in sloping galleries (Georgeson), 484.
George, E. P., Jánossy, L. & McCaig, M. The 'second maximum' of the shower transition curve of cosmic radiation, 219.
Georgeson, E. H. M. The free streaming of gases in sloping galleries, 484.
Goodeve, C. F., Lythgoe, R. J. & Schneider, E. E. The photosensitivity of visual purple solutions and the scotopic sensitivity of the eye in the ultra-violet, 236.
Green, A. E. Stress systems in aeolotropic plates, IV, 173.

Hinshelwood, C. N. *See* Smith & Hinshelwood.
Ho, C. C. *See* Chang & Ho.

Jahn, H. A. Diffuse scattering of X-rays by crystals, II, 476.
Jánossy, L. *See* George, Jánossy & McCaig.
Jeffreys, H. On the significance tests for the introduction of new functions to represent measures, 256.
Jenkins, R. O. *See* Benjamin & Jenkins.

Kothari, D. S. & Singh, B. N. Thermodynamics of a relativistic Fermi-Dirac gas, 414.
Krishnan, R. S. & Nahum, E. A. Cross-section measurements for disintegrations produced by deuterons in the heavy elements, 321.
Krishnan, R. S. & Nahum, E. A. Excitation function measurements for disintegrations produced by deuterons in the heavy elements, 333.

Lattice spacings of solid solutions in magnesium of metals (Raynor), 107.
Lipson, H. *See* Edwards & Lipson.
Lythgoe, R. J. *See* Goodeve, Lythgoe & Schneider.

McCaig, M. *See* George, Jánossy & McCaig.
Macdougall, G. *See* Ockrent & Macdougall.
Molecular structure and rubber-like elasticity (Bunn), 40, 67, 82.

Nahum, E. A. *See* Krishnan & Nahum.
Neighbours, arrangements with given number (Chang & Ho), 345.
Nitrogen, active (Lord Rayleigh), 123, 140.

Ockrent, C. & Macdougall, G. Surface energy in liquid/solid systems, 151.

Peng, H. W. *See* Fuchs & Peng.
Photocombination of hydrogen and chlorine (Ritchie & Taylor), 423.
Polychlorene, polymers (Bunn), 40, 67, 82.
Polychloroprene, structure (Clews), 100.

Quantum mechanics, physical interpretation (Dirac), 1.

Rayleigh, Lord. Further studies on active nitrogen, 123, 140.
Raynor, G. V. The lattice spacings of the primary solid solutions in magnesium of the metals of group IIIB and of tin and lead, 107.
Ritchie, M. & Taylor, D. The photocombination of hydrogen and chlorine in oxygen-free systems, 423.

Sarginson, K. The effect of thermal vibrations on the scattering of X-rays, II, 305.
Schneider, E. E. *See* Goodeve, Lythgoe & Schneider.
Significance tests for the introduction of new functions to represent measures (Jeffreys), 256.
Singh, B. N. *See* Kothari & Singh.
Smith, J. R. E. & Hinshelwood, C. N. The detection and inhibition of free radical chain reactions, 237.
Southwell, R. V. Some practically important stress-systems in solids of revolution, 367.
Stress systems in aeolotropic plates (Green), 173.
Stress-systems in solids of revolution (Southwell), 367.
Surface energy relations in liquid/solid systems (Ockrent & Macdougall), 151.

Taylor, D. *See* Ritchie & Taylor.
Thermodynamics of a relativistic Fermi-Dirac gas (Kothari & Singh), 414.

Visual purple, photosensitivity (Goodeve, Lythgoe & Schneider), 236.

Wilson, A. J. C. Imperfections in the structure of cobalt, 277.

X-rays, effect of thermal vibrations on scattering, II (Sarginson), 305.
X-rays, effect of thermal vibrations on scattering, III (Born), 397.
X-rays, diffuse scattering by crystals (Jahn), 476.

END OF THE ONE HUNDRED AND EIGHTIETH VOLUME (SERIES A)

I. A. R. I. 75.

IMPERIAL AGRICULTURAL RESEARCH
INSTITUTE LIBRARY
NEW DELHI.

Date of issue.	Date of issue.	Date of issue.